Institut für Stochastik und Wirtschaftsmathematik
Operations Research and Control Systems (ORCOS)

(Signature Supervisor)

An Otemachi Approach On Portfolio Optimization

MASTER THESIS

submitted in partial fulfilment of the requirements for the degree of

Master of Science

as part of the studies

Financial and Actuarial Mathematics

written by

Daniel Elmar Toegl
Matriculation number 0730713

at:
Department of Statistics and Mathematical Methods in Economics

Supervisor: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Gernot Tragler
Collaboration: DEPFA Bank plc, Tokyo

Vienna, June 1, 2016

(Signature Author)

Contents

Acknowledgements

I would like to express my deepest gratitude to my advisor, Prof. Tragler. Not only did I learn a lot from him inside and outside his classes, he has always supported me throughout my abstruse idea of collaborating with a Japanese investment bank for this thesis along with his constructive criticism. It is beyond doubt that without the freedom and support he provided in this research project, it would have never been possible to achieve this milestone in my academic and personal development. Living in Tokyo and experiencing the Japanese "Wallstreet" in Otemachi was not only beneficial for my studies, it also enriched my life and will have a lasting impact on my future career.

I am also grateful for Prof. Grandits' time and support ahead of my take-off to Japan. My acknowledgment also goes to all DEPFA Bank staff in Tokyo and most certainly to my risk manager, Linus Lidberg, who enabled me this outstanding opportunity and introduced me to financial markets. Special thanks are given to the International Office who granted me the KUWI research scholarship. This five month research placement in Tokyo would not have been possible without their incredible financial support.

Studying mathematics has never been an easy task for me, but it was one of my best decisions to join the master's degree program at the FAM department at Vienna University of Technology. To name only a few – Arpad Pinter, Prof. Rheinlaender and Prof. Schmock – who opened my mind to new perspectives throughout the course of my studies. Their interest in students, combined with their way of teaching and level of support is incomparable and by far not to be taken for granted.

Last but not the least, I would have never come this far without the inspiring and encouraging words of my friends and family, especially Nicole and Florian who played a major role during the final phase of my studies in Vienna. I owe more than thanks to my parents Angelika and Reinhard who would even visit me just for a weekend in Japan to express their love and support. To me, you are all my globally optimal solution in life. Mina-san, hontou ni arigatou gozaimashita.

Abstract

This thesis discusses several approaches of portfolio optimization in order to better understand modern portfolio theory. It goes beyond the Basel approaches and discusses the necessity of using coherent risk measures. The "Otemachi-Approach" is introduced as the process on how to decide to optimize one's portfolio from a risk manager's view in Japan. It is argued how to hedge earthquake risks within a Japanese stock and fixed income portfolio. In testing this hypothesis, it is proven to be successful by demanding earthquake technology related firms and gum as a commodity to the portfolio. The common idea of hedging by investments in the construction industry is debunked due to back-testing results of Japanese market moves. By using an extremal representation of CVaR, the optimal control problem is reformulated as a bi-level optimization problem. Restrictions of Miller and Yang's model [36] are loosened by using a Record-to-Record algorithm and the usage of the Wasserstein metric to measure the quality of the approximated CVaR from the model is proposed.

Key words: Portfolio optimization, Conditional value-at-risk, Risk measures, Stochastic optimal control, Financial markets

Chapter 1.

Preface

In this work, we imagine being a risk manager of an investment bank in Japan. As Chief Risk Officer it is our job to consider how new ventures fit into the company's portfolio, while taking Japanese market peculiarities into account. Part of our job, is to answer the following questions:

- What is the correlation of the performance of the new venture with the rest of the company's business?
- When the rest of the business is experiencing difficulties, will the new venture provide poor returns, or will it have the effect of dampening the ups and downs in the rest of the business?

Companies must take risks if they are to survive and prosper. The risk management function's primary responsibility is to understand the portfolio of risks that the company is currently taking and the risks it plans to take in the future. It must decide whether the risks are acceptable and, if they are not acceptable, what action should be taken. As all fund managers know, there is a trade-off between risk and return when money is invested. The greater the risk taken, the higher the return that can be realized.

Although there is a difference in the specific definitions of risk and uncertainty, for most of our purposes and in most financial literature the two terms are used interchangeably.

Applied Operations Research and Financial Mathematics provide the essential tools to investigate these topics and to develop new methods, models and algorithms for observing and managing risk in a bank's portfolio. The Japanese market is especially interesting, as it was shattered by several big events which underpinned the importance of risk management and proved, that there are many possible fields of research in Mathematics as well as for economists.

Case studies on Japanese portfolios shall test the performance of the provided tools and will measure, how well they could prepare for surprising events like the Tohoku earthquake, including the Fukushima power plant disaster.

1.1. DEPFA Bank

DEPFA Bank was originally founded in 1922 as a German government agency. In 1951, the bank took the name Deutsche Pfandbriefanstalt, but was being kept under government control, before being privatized in 1991. In order to remove the image of being a German government bank, the name was then changed to DEPFA. At the initial public offering (IPO), the market capitalization of the company was approximately EUR 400 mn. After legal restructuring in the late 90s, the bank moved its headquarters to Dublin, Ireland.

In October 2007, DEPFA Bank was purchased by German Hypo Real Estate, a subsidary of Bayerische Hypo- und Vereinsbank. In 2008, the bank ran into liquidity problems as a result of the economic and financial turmoil caused by Lehmann in the United States and followed by the Irish Banking Crisis until 2010. Through a series of bailouts, the German government ended up with 100% ownership of DEPFA's parent company Hypo Real Estate. Due to EU requirements, DEPFA was about to be privatized again in 2014, but the German government decided to unwind the bank's assets herself. In December 2014, DEPFA's transfer to the German "Bad Bank" FMS Wertmanagement was completed.

DEPFA's Tokyo branch mainly focused on Asian-Pacific markets as well as South America. Infrastructure financing heavily influenced the townscapes of cities like Tokyo and Sao Paulo.

In 2015, the winding down process of DEPFA's Tokyo branch started and will be completed by summer 2016.

1.2. Introduction

The "financial tsunami"[1] in 2008 shattered the industry. Its ramifications are far-reaching and the lessons learned will be embedded in risk management practices for years to come. A central issue in modern risk management is therefore the measurement of risk, as the need to quantify risks arises in many different contexts.

Definition 1 (Risk)**.** *In fact, one way to define risk is the uncertainty of future outcomes. An alternative definition might be the probability of an adverse outcome.*

With this work, we would like to give an overview of existing approaches to measuring risk in financial institutions - such as DEPFA Bank. In discussing strengths and weaknesses of these standard approaches - as they are used in the Basel[2] regulations - we focus on practical aspects. Theoretical properties of the risk measures (coherence) will be discussed later in this work, while taking a Japanese view on markets.

In practice, risk measures are used for a variety of purposes. Among the most important are the following.

- **Determination of risk capital and capital adequacy**

  One of the principal functions of risk management in the financial sector is to determine the amount of capital a financial institution needs to hold as a buffer against unexpected future losses on its portfolio in order to satisfy a regulator, who is concerned with the solvency of the institution. A related problem is the determination of appropriate margin requirements for investors trading at an organized exchange, which is typically done by the clearing house of the exchange.

- **Management tool**

  Risk measures are often used by management as a tool for limiting the amount of risk a unit within a firm may take. For instance, traders in a bank are often constrained by the rule that the daily 95% Value-at-Risk of their position should not exceed a given bound.

[1] Quote of Federal Reserve Chairman Alan Greenspan [44]

[2] The Basel Accords are a global and voluntary regulatory framework on bank capital adequacy, stress testing, and market liquidity risk.

- **Insurance premiums**

  Insurance premiums compensate an insurance company for bearing the risk of the insured claims. The size of this compensation can be viewed as a measure of the risk of these claims.

1.3. Lessons Learned

The following list of Hull [21] sums up some lessons to be learned from the last financial crisis. Especially risk managers should keep an eye on the following points. We will tackle some of these points throughout this work.

- Risk managers should be watching for situations where there is irrational exuberance and make sure that senior management recognize that the good times will not last forever.
- Correlations always increase in stressed markets. In considering how bad things might get, risk managers should not use correlations that are estimated from data collected during normal market conditions.
- Recovery rates decline when default rates increase. This is true for almost all debt instruments, not just mortgages. In considering how bad things might get, risk managers should not use recovery rates that are estimated from data collected during normal market conditions.
- Risk managers should ensure that the incentives of traders and other personnel encourage them to make decisions that are in the interests of the organization they work for. Many financial institutions have revised their compensation policies as a result of the crisis. Bonuses are now often spread out over several years rather than all being paid at once. If good performance in one year is followed by bad performance in the next, part of the bonus for the good-performance year that has not yet been paid may be clawed back.
- If a deal seems too good to be true, it probably is. AAA-rated tranches of structured products promised returns that were higher than the returns promised on AAA bonds by 100 basis points, or more. A sensible conclusion from this for an investor would be that further analysis is needed because there are likely to be risks in the tranches that are not considered by rating agencies.

- Investors should not rely on ratings. They should understand the assumptions made by rating agencies and carry out their own analyses.
- Transparency is important in financial markets. If there is a lack of transparency (as there was for ABS CDOs), markets are liable to dry up when there is negative news.
- Re-securitization, which led to the creation of ABS CDOs and CDOs of CDOs, was a badly flawed idea. The assets used to create ABSs in the first leg of the securitization should be as well diversified as possible. There is then nothing to be gained from further securitization.

DEPFA's Lessons

All financial institutions finance long-term needs with short-term sources of funds to some extent. But a financial institution that relies too heavily on short-term funds is likely to expose itself to unacceptable liquidity risks - like DEPFA Bank.

During the period leading up to the credit crisis of 2007, there was a tendency for subprime mortgages and other long-term assets to be financed by commercial paper while they were in a portfolio waiting to be packaged into structured products.

Conduits and special purpose vehicles had an ongoing requirement for this type of financing. The commercial paper would typically be rolled over every month. For example, the purchasers of commercial paper issued on April 1 would be redeemed with the proceeds of a new commercial paper issued on May 1; this new commercial paper issue would in turn be redeemed with another new commercial paper issued on June 1; and so on. When investors lost confidence in subprime mortgages in August 2007, it became impossible to roll over commercial paper. In many instances, banks had provided guarantees and had to provide financing. This led to a shortage of liquidity. As a result, the credit crisis was more severe than it would have been if longer-term financing had been arranged.

Many of the failures of financial institutions during the crisis (e.g., Lehman Brothers and Northern Rock) were caused by excessive reliance on short-term funding. Once the market (rightly or wrongly) becomes concerned about the health of a financial institution, it can be impossible to roll over the financial institution's short term funding.

The Basel Committee has recognized the importance of liquidity risks by introducing liquidity requirements in Basel III.

Subsequently, in our discussion of portfolio theory, we will consider measures of risk that are used when developing the theory.

Looking at the Japanese market, natural disasters (earthquakes, tsunamis, typhoons) have to be more taken into account than in several other regions of the world. Moreover, most Japanese Government Bonds (JGBs) are held within Japan[3], and should be therefore included to any representative Japanese portfolio. After the Dodd Frank Act (see Acharya's and Richardson's work [2] for further details) was signed in 2010, portfolio optimization will be less favored by banks, but widely used by funds and other organizations, as banks are not supposed to make their profits by managing portfolios, but by trading and its margins.

1.4. Japanese Market Peculiarities

The Lost Decade or Ushinawareta Juunen is the time after the Japanese asset price bubble's collapse within the Japanese economy. The term originally referred to the years from 1991 to 2000, but recently the decade from 2001 to 2010 is often included so that the whole period of the 1990s to the present is referred to as the Lost Two Decades or the Lost 20 Years - Ushinawareta Nijuunen.

Economist Paul Krugman has argued that Japan's lost decade is an example of a liquidity trap (a situation in which monetary policy is unable to lower nominal interest rates because these are close to zero). He explained how truly massive the asset bubble was in Japan by 1990, with a tripling of land and stock market prices during the prosperous 1980s.[4] Back then, the Nikkei hit an intra-day high of 38,957.44 on December 29, 1989 compared to 7,054.98 on March 10, 2009. Japan's high personal savings rates (>90% of JGBs are held within Japan), driven in part by the demographics of an aging population, enabled Japanese firms to rely heavily on traditional bank loans from supporting banking networks, as opposed to issuing stock or bonds via the capital markets to acquire funds.

[3] Statistics can be directly found on the BoJ's website: https://www.boj.or.jp/en/statistics/boj/other/mei/mei.htm/
[4] According to DEPFA colleagues, restaurants topped misou soups with gold in Tokyo.

Lam and Tokuoka [27] explained for the IMF that despite the rise in public debt, Japanese Government Bond (JGB) yields have remained low and stable, supported by steady inflows from the household and corporate sectors, high domestic ownership of JGBs, and safe-haven flows from heightened sovereign risks in Europe. Over time, however, the market's capacity to absorb new debt will likely shrink as population ages and risk appetite recovers. The Financial Times wrote: Foreign ownership of Japanese government bonds has risen to a record JPY 86 tn, or just over 9% of the total outstanding, underscoring the view among crisis-rattled investors that the country's relative stability outweighs the risks posed by its vast public debt.

Comparing this to Austria's national bank's press release from 2011 (ÖNB), stating: "The bulk of Austria's bonds" with more than 80% are held by foreign investors. These numbers shall be seen in context with the Japanese public debt, which exceeded one quadrillion yen (approx. $10.46 tn when published) in 2013, more than twice the annual gross domestic product of Japan.

Aside the financial crisis, Japan has always been heavily shattered by natural disasters. Two out of the five most expensive natural disasters in recent history have occurred in Japan, costing $181 billion in the years 2011 and 1995 only. Japan has also been the site of some of the 10 worst natural disasters of the 21st century. The types of natural disasters in Japan include tsunamis, floods, typhoons, earthquakes, and volcanic eruptions. The country has gone through many years of natural disasters, affecting its economy, development, and social life.[5] This also results in a severe earthquake risk in Japanese portfolios. Regularly, well known Japanese corporates happen to be cause of scandals along their ties with the Yakuza - the Japanese mafia[6]. A popular example is:

- Olympus Corp. used fraudulent takeovers to hide $1.7 billion in losses over 13 years, starting in the 1990s.

The Japanese market is also especially interesting as the economy of Japan is the third-largest in the world by nominal GDP and the fourth-largest by purchasing power parity (PPP) and is the world's second largest developed economy.

[5] Statistics from: https://en.wikipedia.org/wiki/Natural_disasters_in_Japan Last accessed on 24 Apr 2016

[6] A recent list can be found at Bloomberg, last accessed on 24 Apr 2016: http://www.bloomberg.com/news/articles/2015-07-21/toshiba-s-accounting-scandal-ranks-among-japan-s-largest-cases

Remark 1 (Quantitative Easing in Japan)**.** *In terms of the BoJ's market share in the JGB market, it renewed its new peak. In May 2015, the BoJ owned 28.2% of the JGB market, measured in value, and 24% measured in aggregate duration risk.*

Chapter 2.

Portfolio Theory

Portfolio theory deals with the problem of constructing – for a given collection of assets – an investment with desirable features. A variety of different asset characteristics can be take into consideration, such as the amount of value, on average, an asset returns over a period of time or the riskiness of reaping returns comparable to the average. The financial objectives of the investor and the tolerance of risk determine what type of portfolios are to be considered desirable. In particular, we shall now use the methods of constrained optimization to construct portfolios for a given collection of assets with desirable features.

Definition 2 (Portfolio)**.** *A grouping of financial assets such as stocks, bonds and cash equivalents, as well as their mutual, exchange-traded and closed-fund counterparts. Portfolios are held directly by investors and/or managed by financial professionals. It can be seen as an investment made in n assets using some amount of wealth W.*

Prudence suggests that investors should construct an investment portfolio in accordance with risk tolerance and investing objectives. Let's think of an investment portfolio as a pie that is divided into pieces of varying sizes representing a variety of asset classes and/or types of investments to accomplish an appropriate risk-return portfolio allocation.

Common integral parts of portfolios are the following and previously mentioned financial assets.

Definition 3 (Stocks)**.** *A stock - also known as shares or equity – is a type of security that signifies ownership in a corporation and represents a claim on part of the corporation's assets and earnings.*

Remark 2. *Stocks are the foundation of nearly every portfolio. Historically, they have outperformed most other investments over the long run.*

Definition 4 (Bonds)**.** *A bond is a debt investment in which an investor loans money to an entity (typically corporate or governmental) which borrows the funds for a defined period of time at a variable or fixed interest rate. Bonds are used by companies, municipalities, states and sovereign governments to raise money and finance a variety of projects and activities. Owners of bonds are debt holders, or creditors, of the issuer.*

Zero-Coupon-Bonds (ZCB) of longer maturities are relatively rare in practice, but still relevant, since many fixed income instruments such as coupon bonds or standard swaps can be viewed as portfolios of Zero-Coupon-Bonds. We follow the standard convention as in McNeil et al. [32] *and normalize the face value at (maturity date T) $p(T,T)$ on the bond to* 1*. The continuously compounded yield of a Zero-Coupon-Bond is defined as $y(s,T) := -\frac{1}{T-s} \ln p(s,T)$, i.e., we have*

$$p(s,T) = \exp(-(T-s)y(s,T)). \tag{2.1}$$

The mapping $T \to y(s,T)$ is referred to as the continuously compounded yield curve at time s. In a detailed analysis of the change in the value of a bond portfolio one takes all yields $y(s,T_i), 1 \leq i \leq d$ as risk factors for d default-free Zero-Coupon-Bonds with maturity T_i.

Remark 3 (JGBs and Samurai Bonds)**.** *Japanese Government Bonds are commonly abbreviated as JGBs and play a massive role in the global bond market. Some foreign issuer bonds are called by nicknames, such as the "Samurai Bond". These can be issued by foreign issuers looking to diversify their investor base away from domestic markets. These bond issues are generally governed by the law of the market of issuance, e.g., a Samurai Bond, issued by an investor based in Europe, will be governed by Japanese law.*

Remark 4. *Bonds are commonly referred to as fixed-income securities and are one of the three main generic asset classes, along with stocks and cash equivalents.*

Definition 5 (Cash Equivalents)**.** *Cash equivalents are assets that are readily convertible into cash, such as money market holdings, short-term government bonds or treasury bills, marketable securities and commercial paper. Cash equivalents are distinguished from other investments through their short-term existence; they mature within 3 months,* *whereas short-term investments are 12 months or less, and long-term investments are any investments that mature in excess of 12 months.*

Risk aversion is a concept in economics and finance, based on the behavior of humans (especially consumers and investors) while exposed to uncertainty to attempt to reduce that uncertainty.

Definition 6 ((Absolute Pratt-Arrows)-Risk Aversion[39]). *Let $u(x)$ be a utility function. Then, the function*

$$A(x) := \frac{-u''(x)}{u'(x)} \tag{2.2}$$

will be interpreted in various ways as a measure of local risk aversion (often referred to "risk aversion in the small"). Individuals may have different risk attitudes. A person is said to be:

- *risk-averse (or risk-avoiding), if he or she would accept a certain payment (certainty equivalent) of less than JPY 50 (for example, JPY 40), rather than taking the gamble and possibly receiving nothing.*
- *risk-neutral, if he or she is indifferent between the bet and a certain JPY 50 payment.*
- *risk-loving (or risk-seeking), if he or she would accept the bet even when the guaranteed payment is more than JPY 50 (for example, JPY 60).*

We expect to be risk averse, especially us, as imaginary Risk Managers. If one can choose between two investments with the same expected return, he or she will choose the less risky one. The opposite are therefore risk loving investors, who would strongly consider taking more risk, to reach some certain amount of, let's say, money.

Remark 5. *In modern portfolio theory, risk aversion is measured as the additional marginal reward an investor requires to accept additional risk. In modern portfolio theory, risk is being measured as standard deviation of the return on investment, i.e., the square root of its variance. In advanced portfolio theory, different kinds of risk are taken into consideration. They are being measured as the n-th radical of the n-th central moment. The symbol used for risk aversion generally is A or A_n.*

2.1. Quantifying Risk

A convenient measure that is often used is the standard deviation of the return over one year. In 1952, Markowitz published his paper on Portfolio Theory and won the Nobel prize for his work in 1990. His mean-variance analysis has been widely used ever since. This thesis, will explore the basic theory of Markowitz and go beyond it. Main results from Markowitz are taken from his works in 1952, 1959 and 1968 in [29], [30] and [31].

Markowitz showed that the variance of the rate of return was a meaningful measure of portfolio risk under a reasonable set of assumptions and he derived the formula for computing the variance of a portfolio. This formula for the variance of a portfolio not only indicated the importance of diversifying your investments to reduce the total risk of a portfolio, but showed how to effectively diversify. The simple Markowitz model is based on several assumptions regarding investor behavior:

1. Investors consider each investment alternative as being represented by a probability distribution of expected returns over some holding period.
2. Investors maximize one-period expected utility, and their utility curves demonstrate diminishing marginal utility wealth.
3. Investors estimate the risk of the portfolio on the basis of the variability of expected returns.
4. Investors base decision solely on expected return and risk, so their utility curves are a function of expected return and the expected variance (or standard deviation) of returns only.
5. For a given risk, investors prefer higher returns to lower returns. Similarly, for a given level of expected return, investors prefer less risk to more risk (risk averse).

Under these assumptions, a single asset or portfolio of assets is considered to be efficient if no other asset or portfolio of assets offers higher expected return with the same (or lower) risk, or lower risk with the same (or higher) expected return.

2.1.1. Measures of Risk Aversion under Expected Utility

Let's take a deeper look into the above mentioned utility curves.

We recall the three attributes of being

- risk-averse (accepting a certain payment rather than taking a gamble and possibly receiving nothing),
- risk-neutral (indifferent between betting and a fixed payment) and
- risk-seeking (preferring the bet rather than a guaranteed payment).

Definition 7 (Utility)**.** *In economics and finance, utility is a measure of preference over some set of goods or results. Utility is usually applied in such constructs as a monotone and concave/convex curve.*

Example 1 (Arnold Schwarzenegger)**.** *The idea of utility can be applied to many other fields. Just think of being rich or poor or about being happy in that context. At some point, it does not really matter any more if one is super rich and gets another man yen[1], whereas on the other end, it could help someone poor to survive.*

"Money doesn't make you happy. I now have $ 50 million but I was just as happy when I had $ 48 million."[2] "

Remark 6 (Equivalence of Risk Aversion and Concavity)**.** *With Jensen's inequality, we can conclude that risk aversion is equivalent to concavity of a utility function. For the basic idea, we can keep in mind the following:*

strict concavity ⇔ strict risk aversion
linearity ⇔ risk neutral
strict convexity ⇔ risk-seeking

Weaker requirements are possible and we will go more into detail with quasi-convexity etcetera when needed.

The higher the curvature of our risk function u, the higher the risk aversion. Let's take a look now at commonly used measures. We generally differentiate between absolute and relative measures. We will also take a deeper look on applications on portfolio measures.

[1] In Japan, a different way of counting is used. When numbers reach 10000 they are counted in groups of ten thousand, which is man. The extra zeros usually are not written, neither.

[2] Arnold Schwarzenegger, Quote retrieved March 1, 2016, from http://www.brainyquote.com/quotes/quotes/a/arnoldschw146572.html

Definition 8 (Arrow-Pratt measures)**.** *We recall the Arrow-Pratt measure formula 2.2 which is as well often referred to as the coefficient of absolute risk aversion. Subsequently, we can define the Arrow-Pratt measure of relative risk aversion, by*

$$RRA(c) = cARA(c) = -\frac{cu''(c)}{u'(c)} \tag{2.3}$$

A current problem Japan is facing in its deflationary environment can be described by risk aversion. In monetary economics, according to Benchimal and Fourcans [9], an increase in relative risk aversion increases the impact of households' money holdings on the overall economy. In other words, the more the relative risk aversion increases, the more money demand shocks will impact the economy. As the BoJ experienced negative yields on short to midterm JGBs by 2016, the national bank now actually gets paid to borrow money – a crucial sign of risk aversion in markets. [3]

In modern portfolio theory, risk aversion is measured as the additional marginal reward an investor requires to accept additional risk. In former approaches, risk is being measured as standard deviation of the return on investment, i.e., the square root of its variance. In advanced portfolio theory, n different kinds of risk are taken into consideration. They are being measured as the n-th radical of the n-th central moment.

Definition 9 (Risk Aversion in Advanced Portfolio Theory)**.** *For the overall portfolio, the symbol A is often used as*

$$A = \frac{\mathbb{E}[r])}{d\sigma}$$

with along with the n-th radical of the n-th central moment μ_n that is used as follows:

$$A_n = \frac{d\mathbb{E}[r])}{d\sqrt[n]{\mu_n}} = \frac{1}{n}\frac{d\mathbb{E}[r]}{d\mu_n} \tag{2.4}$$

to describe risk aversion in advanced portfolio theory with the derived expected utility $\mathbb{E}$ *of the additional reward r.*

2.1.2. Approaches to Managing Risk

The following subsections follow Hull's work in 2012 on risk management [21].

[3]See Reuters: http://www.reuters.com/article/us-japan-bonds-idUSKCN0W34H3, last accessed on 23-Apr-2016.

In general, we can say that there are two broad risk management strategies open to the bank, any fund or any other organization. One is to identify risks one by one and to handle each one separately, which is often referred to as risk decomposition. The other one is to reduce risk by being well diversified, which is also known as risk aggregation. In practice, both approaches are used.

Market Risks

Market risks arise primarily from trading activities. One might most likely be exposed to interest and exchange rates, equity prices and other market variables. These risks are usually managed directly by traders, who, at the end of each day, ensure that risk limits are not exceeded. Risk managers then aggregate the residual risks from activities of all traders to determine the total risk faced by the bank from movements in market variables – the risk factors. A well diversified portfolio reduces exposure to market movements, as moves are partly setting off each other. If risks are unacceptably too high, the reason has to be determined and corrective (hedging) action must be taken.

Credit Risks

Credit risks are managed traditionally by ensuring that the credit portfolio is well diversified. Only lending money to a single borrower is naturally a huge risk, whereas it is unlikely that 100 borrowers in different markets will default simultaneously. Nevertheless, diversification reduces non-systematic risk and does not eliminate systematic risk. That means, one always faces the risk of economic downturns and a resulting increase of the probability of default by borrowers, or huge losses on the market risk side. Diversification also does not protect an international bank, which diversified worldwide across industries, a global economic downturn. With the advent of credit derivatives, a risk decomposition approach can be used to handle credit risk one by one.

2.2. Building a Portfolio

Building a portfolio needs a lot of thought and requires certain strategies. In general, one can decide between five main categories of portfolio[4]:

1. **Aggressive**: The strategy emphasizes capital appreciation as a primary investment objective. It is especially suitable for young adults because their lengthy investment horizon enables them to ride out market fluctuations better than investors with a short-term investment horizon.
2. **Defensive**:Such a strategy entails regular portfolio rebalancing to maintain one's intended asset allocation; buying high-quality, short-maturity bonds and blue-chip stocks; diversifying across both sectors and countries; placing stop loss orders; and holding cash and cash equivalents in down markets. Such strategies are meant to protect investors against significant losses from major market downturns.
3. **Fixed-Income**: The main goal of this strategy is the, preservation of capital with a guaranteed income stream. This is especially of importance when one is approaching retirement.
4. **Speculative**: Speculatie strategies are often called "pure gamble". These "plays" could be initial public offerings (IPOs) or stocks that are rumored to be takeover targets. Technology or health care firms that are in the process of researching a breakthrough product, or a junior oil company which is about to release its initial production results, would also fall into this category. Speculative stocks are typically trades, and not your classic *buy and hold* investment.
5. **Hybrid**: Basically, a hybrid portfolio would include a mix of stocks and bonds in a relatively fixed allocation proportions. This type of approach offers diversification benefits across multiple asset classes as equities and fixed income securities tend to have a negative correlation with one another.

Imagine being young: One might rather build up a more aggressive portfolio, which will be riskier, as we have less to lose and are eager to build up a certain amount of money for the future. Thinking of one's own pension savings, we can apply this thought as well. Shortly ahead of retirement, we do not want to lose our savings anymore and

[4]These defintions mainly follow http://www.investopedia.com/terms/d/defensiveinvestmentstrategy.asp, last accessed on 29 May 2016

rather stick to saver risk allocation as we will have to live from those savings in the near future. Beside diversification, one should think of a time horizon, markets, equities and possible other risks.

Additionally, the common problem of traders and risk managers leads to different approaches on risk management and the ability to find a compromise. In general, one could say that if traders were following approaches from risk management, the could not make any deals as they pricing is usually too far away from real market prices.

2.2.1. Trader's Risk Management – A Business Snapshot

Factor-sensitivity measures give the change in portfolio value for a given predetermined change in one of the underlying risk factors: Typically they take the form of a derivative (in the calculus sense). Important factor sensitivity measures are the duration for bond portfolios and the Greeks for portfolios of derivatives. While these measures provide useful information about the robustness of the portfolio value with respect to certain well-defined events, they cannot measure the overall riskiness of a position. Moreover, factor-sensitivity measures create problems in the aggregation of risks.

What are "Greeks"?

Greeks are dimensions of risk involved in taking a position in an option (or other derivative). Each risk variable is a result of an imperfect assumption or relationship of the option with another underlying variable. Various sophisticated hedging strategies are used to neutralize or decrease the effects of each variable of risk.

Neutralizing the effect of each variable requires substantial buying and selling and, as a result of such high transactions costs, many traders only make periodic attempts to re-balance their options portfolios. With the exception of *vega* (which is not a Greek letter), each measure of risk is represented by a different letter of the Greek alphabet:

Definition 10 (Greeks)**.**

Δ (Delta) represents the rate of change between the option's price and the underlying asset's price - in other words, price sensitivity.

Θ *(Theta) represents the rate of change between an option portfolio and time, or time sensitivity.*

Γ *(Gamma) represents the rate of change between an option portfolio's delta and the underlying asset's price - in other words, second-order time price sensitivity.*

Vega represents the rate of change between an option portfolio's value and the underlying asset's volatility - in other words, sensitivity to volatility.

ρ *(Rho) represents the rate of change between an option portfolio's value and the interest rate, or sensitivity to the interest rate.*

For a given portfolio it is not possible to aggregate the sensitivity with respect to changes in different risk factors. For instance, it makes no sense to simply add the delta and the vega of a portfolio of options. Factor-sensitivity measures cannot be aggregated across markets to create a picture of the overall riskiness of the portfolio of a financial institution.

Dynamic hedging in practice generally relies on using the commonly referred to the Greeks. This chapter only offers a mathematical view on the Greeks and follows Hull [21] to some extent. A detailed interpretation can also be found in Hull. Each of the Greeks measures a different aspect of the risk in a trading position.

The trading function within a financial institution is referred to as the front office; the part of the financial institution that is concerned with the overall level of the risks being taken, capital adequacy, and regulatory compliance is referred to as the middle office; the record keeping function is referred to as the back office.

In a typical arrangement at a financial institution, the responsibility for a portfolio of derivatives dependent on a particular underlying asset is assigned to one trader or to a group of traders working together. For example, one trader at DEPFA Bank might be assigned responsibility for all derivatives dependent on the value of the Australian dollar. A computer system calculates the value of the portfolio and Greek letters for the portfolio. Limits are defined for each Greek letter, and special permission is required if a trader wants to exceed a limit at the end of a trading day. The delta limit is often expressed as the equivalent maximum position in the underlying asset. For example, the Δ limit of DEPFA Bank on a stock might be specified as \$10 million. If the stock price is \$50, this means that the absolute value of Δ as we have calculated it can be no more that 200,000.

The vega limit is usually expressed as a maximum dollar exposure per 1% change in the volatility. As a matter of course, options traders make themselves Δ neutral – or close to Δ neutral – at the end of each day. Γ and vega are monitored, but are not usually managed on a daily basis. Financial institutions often find that their business with clients involves writing options and that as a result they accumulate negative Γ and vega. They are then always looking out for opportunities to manage their gamma and vega risks by buying options at competitive prices.

There is one aspect of an options portfolio that mitigates problems of managing gamma and vega somewhat: Options are often close to the money when they are first sold so that they have relatively high Γs and vegas. When some time has elapsed, however the underlying asset price has often changed sufficiently for them to become deep-out-of-the-money or deep-in-the-money. Their Γs and vegas are then very small and of little consequence. The nightmare scenario for an options trader is where written options remain very close to the money as the maturity date is approached.

2.2.2. Diversification

A common technique is to invest in a variety of assets, following the saying:

"Don't put all of your eggs in one basket."

Definition 11 (Diversification)**.** *The concept of diversification is to allocate capital in a way (for example by multiple investments) in order to reduce exposure to an asset or risk.*

Best Performing Portfolio as a Benchmark

This chapter follows the work of Platen [38]. The main result of his work can be summarized as follows:

Better performance through better diversification!

We will come accross the idea of diversification throughout this work. Our main concern will be a problem of dynamic stochastic optimization: to find a trading strategy whose wealth appears "better" when compared to the wealth generated by any other strategy, in the sense that the ratio of the two processes is a super-martingale. If such a strategy exists, it is essentially unique and is called numéraire portfolio.

Theorem 1 (Diversification Theorem - Platen)**.** *In a regular market, any diversified portfolio is an approximate numéraire portfolio.*

In a theoretical perfectly diversified portfolio, Platen was able to show that diversification outperforms every global index even in times of crisis.

2.2.3. Rates of Return of Assets and Portfolios

To build up a small example portfolio with equities (company shares) we shall concern two basic features of an asset. One is the average return of an asset over a period of time. The second characteristics is how risky it is to obtain similar returns comparable to the average over the investment period. This chapter follows Atzberger's work on Portfolio Theory [6].

Definition 12 (Rate of Return)**.** *For an asset i with value $S(0)$ at time 0 and value $S(T)$ at time T, the rate of return ρ_i is defined by*

$$S(T) = (1+\rho_i)S(0). \tag{2.5}$$

The rate of return of an asset is also sometimes referred to as the yield of the asset.

The rate of return can be thought of as an effective interest rate, which would be required for a deposit of $S(0)_i$ into a savings account at a bank to obtain the same change in value as the asset over the period $[0,T]$ for the i^{th} asset of a given collection of n assets $S_1, S_2, \ldots, S_n$. Since the outcome of an investment in an asset has some level of uncertainty, the value $S(T)_i$ is unknown at time 0. To model this uncertainty or risk, we consider $S(T)_i$ as a random variable. Therefore, ρ_i is also a random variable.

Definition 13. *To characterize the asset, we shall consider the average rate of return, defined by using the expectation as:*

$$\mu = E[\rho], \tag{2.6}$$

which is also referred as the expected rate of return, cf. Figure 2.2

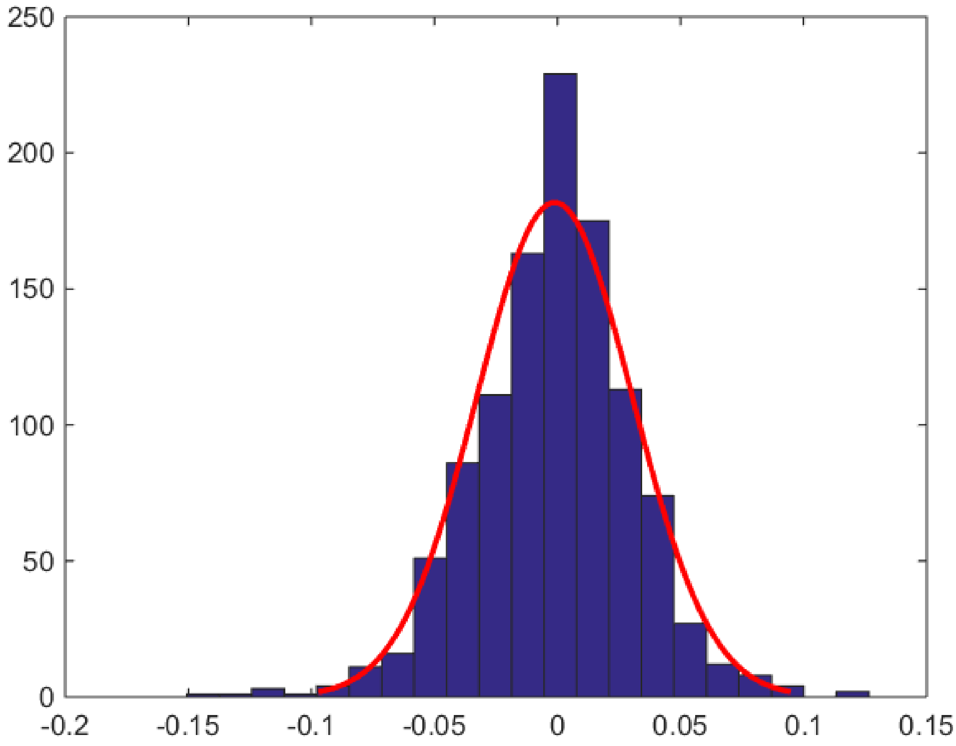

Figure 2.1.: Portfolio Returns vs. Normal Distribution

In Figure 2.2 we compute the portfolio returns and map it against their respective normal distribution. Ye can clearly see, that these portfolio returns are not normally distributed. This fact needs to be kept in mind for all further models and calculations.

While the expected rate of return characterizes an asset and gives indication of how large the returns may be, it does not capture the uncertainty in obtaining a comparable return rate to the average.

To quantify how much the rate of return deviates from the expected return and in order to capture the riskiness, we shall use the variance defined by

$$\sigma_i^2 = Var[\rho] = E[|\rho - \mu|^2]. \tag{2.7}$$

It is important not only to take individual returns when choosing an investment, but also how the returns are coupled with each other. A basic investment strategy to reduce risk in which an asset loses value should a given event occur, is to try to find another asset which increases in value should this event materialize. This requires that the assets exhibit a coupling in which values move in opposite directions should the event occur. To quantify this for random rates of return, we shall use the covariance for the returns, defined by

$$\sigma_{i,j} = E\left((\rho_i - \mu_i)(\rho_j - \mu_j)\right). \tag{2.8}$$

We see that the covariance is symmetric, $\sigma_{i,j} = \sigma_{j,i}$ and for $i = j$ it follows that $\sigma_{i,i} = \sigma_i^2$. The coupling of all n assets can be summarized in the covariance matrix

$$V = \begin{pmatrix} \sigma_1^2 & \cdots & \sigma_{1,n} \\ & \ddots & \\ \sigma_{n,1} & \cdots & \sigma_n^2 \end{pmatrix}. \tag{2.9}$$

Proposition 1 (Covariance Matrix)**.** *The covariance matrix V is positive semi-definite and hence correlation matrices are positive semi-definite.*

Covariance Matrix is positive semi-definite. A proof can be found in McNeil et al. [32] on page 64. □

In our setting, we are using market data. We can therefore expect V to be positive definite as we would expect market changes everyday and therefore the determinant is expected to be non-zero, as its Eigenvalues will be non-zero. V will therefore be invertible.

Remark 7 (Risk Factors at DEPFA)**.** *A common practice at DEPFA was to monitor all kinds of risk factors. Sometimes, the head office in Dublin decided to remove risk factors from the system, or some risk factors actually stalled as the portfolio was not affected by them anymore. It was therefore necessary to identify these risk factors and exclude them from the Covariance matrix to use a positive definite covariance matrix to calculate the daily risk profile.*

As mentioned above, a portfolio is an investment made in n assets using some amount of wealth W. Let W_i denote the amount of money invested in the i^{th} asset. We shall allow negative values for W_i, which can be interpreted as short selling an asset. Short selling means, the sale of a security that is not owned by the seller, or that the seller has borrowed. Short selling is usually motivated by the belief that prices decline, which enables them to be bought back at a lower price and therefore earn profit. It is used either for speculative wins, or for hedging downside risk of a long position.

As we want to invest the total amount of our wealth, we get

$$W = \sum_{i=1}^{n} W_i, \tag{2.10}$$

From now on, it will be easier to describe our investments in terms of relative values, such as the rate of return and the relative portion of wealth invested in an asset. We therefore define

$$w_i = \frac{W_i}{W}, \tag{2.11}$$

which implies by equation 2.10, that $\sum_{i=1}^{n} w_i = 1$.

Definition 14 (Value of Portfolio). *The value Q_p of a portfolio p at the time t can be expressed as*

$$Q_p(t) = \sum_{i=1}^{n} \frac{W_i}{S_i(0)} S_i(t), \tag{2.12}$$

where $Q_p(0) = W$.

Definition 15 (Rate of Return for a Portfolio). *The rate of return ρ_p of the portfolio p at time t is given by:*

$$\rho_p(t) = \frac{Q_p(t) - Q_p(0)}{Q_p(0)}. \tag{2.13}$$

Using the above definitions, we can easily conclude the following Lemma:

Lemma 1 (Rate of Return for a Portfolio). *The rate of return for a portfolio is the weighted average of the rates of return of the assets, where the weights are determined by the proportion of the wealth invested in each asset.*

Proof.

$$\rho_p(t) = \frac{Q_p(t) - Q_p(0)}{Q_p(0)}$$

As $Q_p(0) = W$, it follows that

$$\rho_p(t) = \frac{\sum_{i=1}^{n} \frac{W_i}{S_i(0)} S_i(t) - W}{W}$$

$$= \sum_{i=1}^{n} \frac{W_i S_i(t)}{W S_i(0)} - \sum_{i=1}^{n} \frac{W_i}{W}$$

which, by using the definition of w_i, is equal to

$$\sum_{i=1}^{n} w_i \frac{(S_i(t) - S_i(0))}{S_i(0)} = \sum_{i=1}^{n} w_i \rho_i$$

□

2.3. Basic Concepts in Risk Management

In this section we discuss essential concepts in quantitative risk management. We need a probabilistic framework for modeling financial risk and give formal definitions for notions such as risk, profit and loss, risk factors and mapping.

2.3.1. General Definitions

Example 2 (Japanese Stock Portfolio)**.** *Consider a fixed portfolio of d stocks and denote by $\lambda_{t,i}$ the number of shares of stock i at time t. The price process of stock i is denoted by $(S_{t,i})_{t\in\mathbb{N}}$. Following standard practice in finance and risk management and according to McNeil et al.* [32]*, we use logarithmic prices as risk factors, i.e., we take $Z_{t,i} := \ln S_{t,i}$, $1 \leq i \leq d$. The risk-factor changes $X_{t+1,i} := \ln S_{t+1,i} - \ln S_{t,i}$ then correspond to the log-returns of the stocks in the portfolio. We get $V_t = \sum_{i=1}^{d} \lambda_i \exp(Z_{t,i})$ for the portfolio value at time t and hence define the portfolio loss L_{t+1} as*

$$L_{t+1} = -(V_{t+1} - V_t) = -\sum_{i=1}^{d} \lambda_i S_{t,i}(\exp(X_{t+1,i}) - 1)).$$

Hence, the linearized loss $L^{\triangle}_{t+1}$ is given by

$$L^{\triangle}_{t+1} = -\sum_{i=1}^{d} \lambda_i S_{t,i} X_{t+1,i}.$$

Given the mean vector and covariance matrix of the distribution of the risk-factor changes it is easy to compute the first two moments of the conditional distribution with *mean vector μ and covariance matrix V. Applied to the mean vector μ_t and covariance V_t of the conditional distribution $F_{X_{t+1}|\mathbf{F}_t}$ of the risk factor changes yields the first two moments of the unconditional distribution F_X of the risk factor changes.*

Example 3 (DEPFA: Stylized Portfolio of Risky Loans)**.** *The German "Pfandbriefe"* [5] *represented the earliest example of covered bonds – that is, bonds backed by long-term assets as collateral – in Europe. Invented in the late 18th century, the Pfandbriefe were traditionally used in two ways: on the one hand, for funding property development by securing loans with mortgages, and on the other, for financing loans to the public sector. The new ownership led to the bank's adopting a new name, Deutsche Pfandbriefanstalt, or DEPFA. Under its new structure, DEPFA operated on a mostly nonprofit basis, with a mandate to grant loans to the residential property market.*

Remark 8. *A loan portfolio is subject to many risks. The most important ones are default risks, i.e., the risk that some counterparties cannot repay their loans, interest-rate risk, i.e., the risk that the present value of the future cash flows from the portfolio is diminished due to rising interest rates; and finally, the risk of losses caused by rising credit spreads.*

We consider a portfolio of loans to m different counterparties, the size of the exposure to counterparty i is denoted by e_i. Following standard practice in credit risk management, the risk-management horizon Δ (not to be mistaken with our "greek" Δ) is taken to be one year so that there is no need to distinguish between two timescales (i.e., t and s). For simplicity, we assume that all loans are repaid at the same date $T > t$ and that there are no payments prior to T. We introduce an risk value $Y_{t,i}$ that represents the default state of counterparty i at time t. We set $Y_{t,i} = 1$ if counterparty i defaults within time period $[0,t]$ and $Y_{t,i} = 0$ otherwise. Again for simplicity we assume a recovery rate of zero, i.e., we assume that upon default obliger i the whole exposure e_i is lost.

In valuing a risky loan we have to take into account the possibility of default. Typically, this is done by discounting the cash flow e_i at a higher rate than the yield $y(t,T)$ of a default-free zero-coupon bond. More precisely, we model the value at time t of such a loan as

$$[ValueofaRiskyLoan] = \exp(-(T-t)(y(t,T) + c_i(t,T)))e_i \tag{2.14}$$

with $c_i(t,T)$ referred to as the credit spread of company i corresponding to the maturity T.

[5] Definitions can be found on: FundingUniverse.com or McNeil et al. [32]

Remark 9 (Credit Spreads)**.** *Credit spreads are often determined from the prices of traded corporate bonds issued by companies with a similar credit quality to the counterparty under consideration. Alternatively, for instance a formal pricing model can be used, which utilizes the market value of the counterparty's stock as main input.*

For simplicity, we ignore variations in credit quality and assume that $c_i(t,T) = c(t,T)$ for all i. Under all these simplifying assumptions the value at time t of our loan portfolio equals

$$[ValueofLoanPortfolio] V_t = \sum_{i=1}^{m} (1 - Y_{t,i}) \exp(-(T-t)(y(t,T) + c(t,T))) e_i. \quad (2.15)$$

This suggests the following $(m+2)$-dimensional random vector of risk factors

$$Z_t = (Y_{t,1}, \ldots, Y_{t,m}, y(t,T), c(t,T))^t. \quad (2.16)$$

L_{t+1} and l_{t+1} are now easy to compute. Due to the discrete nature of the default indicators and the long time horizon, linearized losses are of little importance in credit risk management. It is apparent from the formulas above in this example that the main difficulty in modeling the loss distribution of loan portfolios is in finding and calibrating a good model for the joint distribution of the default indicators $Y_{t+1,i}$, $1 \leq i \leq m$.

2.3.2. Approaches to Risk Measurement

A central issue in modern risk management is the measurement of risk, as the need to quantify risks arises in many different contexts.

Existing approaches to measuring the risk of a financial position can be grouped into four different categories according to McNeil et al. [32]:

1. The Notional-Amount approach
2. Factor-Sensitivity measures
3. Risk Measures based on the loss distribution
4. Risk Measures based on scenarios

Notional-amount approach

This is the oldest approach to quantifying the risk of a portfolio of risky assets. In the notional-amount approach, the risk of a portfolio is defined as the sum of the notional values of the individual securities in the portfolio, where each notional value may be weighted by a factor representing an assessment of the riskiness of the broad asset class to which the security belongs. Variants of this approach are still in use in the standardized approach of the Basel Committee rules on banking regulation. See Crouhy, Galai and Mark [15] for further details.

The advantage of the notional-amount approach is its apparent simplicity. However, from an economic viewpoint the approach is flawed for a number of reasons. To begin with, the approach does not differentiate between long and short positions and there is no netting.

Remark 10 (Long and Short Positions)**.** *In finance, a long position in a security, such as a stock or a bond, or equivalently to be long in a security, means the holder of the position owns the security and will profit if the price of the security goes up. "Going long" is the more conventional practice of investing and is contrasted with going short. An options investor goes long on the underlying instrument by buying call options or writing put options on it.*

In contrast, a short position in a futures contract or similar derivative means that the holder of the position will profit if the price of the futures contract or derivative goes down.

Example 4 (No Netting)**.** *The risk of a long position in a foreign currency hedged by an offsetting short position in a currency forward would be counted as twice the risk of the unhegded currency position.*

Moreover, the approach does not reflect the benefits of diversification on the overall risk of the portfolio.

Example 5 (No diversification benefits)**.** *If we use the notional-amount approach, it appears that a well-diversified credit portfolio consisting of loans to m companies that default more or less independently has the same risk as a portfolio where the whole amount is lent to a single company.*

Finally, the notional-amount approach has problems in dealing with portfolios of derivatives, where the notional amount of the underlying and the economic value of the derivative position can differ widely.

Factor-Sensitivity Measures

Factor-sensitivity measures give the change in portfolio value for a given predetermined change in one of the underlying risk factors; typically they take the form of a derivative (in the calculus sense). Important factor-sensitivity measures are the duration for bond portfolios and the so called Greeks for portfolios of derivatives.

Remark 11 (Greeks)**.** *In mathematical finance, the Greeks are the quantities representing the sensitivity of the price of derivatives such as options to a change in underlying parameters on which the value of an instrument or portfolio of financial instruments is dependent. See 2.2.1.*

While these measures provide useful information about the robustness of the portfolio value with respect to certain well-defined events, they cannot measure the overall riskiness of a position. Moreover, factor-sensitivity measures create problems in the aggregation of risks.

1. For a given portfolio it is not possible to aggregate the sensitivity with respect to changes in different risk factors. For instance, it makes no sense to simply add the Δ and the vega of a portfolio of options.
2. Factor-sensitivity measures cannot be aggregated across markets to create a picture of the overall riskiness of the portfolio of a financial institution.

Hence these measures are not very useful for capital-adequacy decisions, but in conjunction with other measures, they can be useful for setting position limits, as DEPFA Bank does.

Risk measures based on loss distributions.

Most modern measures of the risk in a portfolio are statistical quantities describing the conditional or unconditional loss distribution of the portfolio over some predetermined horizon Δ. Examples include the variance, the Value-at-Risk and the expected shortfall,

which we discuss in more detail later in this work.
It is of course problematic to rely on any one particular statistic to summarize the risk contained in a distribution. However, the view that the loss distribution as a whole gives an accurate picture of the risk in a portfolio has much to commend it:

1. losses are the central object of interest in risk management and so it is in natural to base a measure of risk on their distribution,
2. the concept of a loss distribution makes sense on all levels of aggregation from a portfolio consisting of a single instrument to the overall position of a financial institution,
3. if estimated properly, the loss distribution reflects netting and diversification effects; and
4. loss distributions can be compared across portfolios.

For instance, it makes perfect sense to compare the loss distribution of a book of fixed-income instruments and of a portfolio of equity derivatives, at least if the time horizon Δ is the same in both cases.

Nonetheless, McNeil et al. [32] described two major problems when working with loss distributions. First, any estimate of the loss distribution is based on past data. If the laws governing financial markets change, these past data are of limited use in predicting future risk. The second, related problem is practical. Even in a stationary environment it is difficult to estimate the loss distribution accurately, particularly for large portfolios, and many seemingly sophisticated risk-management systems are based on relatively crude statistical models for the loss distribution (incorporating, for example, untenable assumptions of normality).

However, this is not an argument against loss distributions. Rather, it calls for improvements in the way loss distributions are estimated and for prudence in the practical application of risk-management models which are based on estimated loss distributions. In particular, risk measures based on the loss distribution should be complemented by information from hypothetical scenarios. Moreover, forward-looking information reflecting the expectations of market participants, such as implied volatilities, should be used in conjunction with statistical estimates (which are necessarily based on historic data) in calibrating models of the loss distribution,

Remark 12 (Implied Volatility)**.** *Implied volatility, a forward-looking and subjective measure, differs from historical volatility because the latter is calculated from known past returns of a security. The implied volatility of an option is a more useful measure of the option's relative value than its price. The reason is that the price of an option depends most directly on the price of its underlying asset. If an option is held as part of a delta neutral portfolio (that is, a portfolio that is hedged against small moves in the underlying's price), then the next most important factor in determining the value of the option will be its implied volatility.*

Implied volatility is so important that options are often quoted in terms of volatility rather than price, particularly between professional traders.

Example 6 (Implied Volatility)**.** *Consider a call option trading at JPY 1.50 with the underlying trading at JPY 42.05. The implied volatility of the option is determined to be 18.0%. A short time later, the option is trading at JPY 2.10 with the underlying at JPY 43.34, yielding an implied volatility of 17.2%. Even though the option's price is higher at the second measurement, it is still considered cheaper based on volatility. The reason is that the underlying needed to hedge the call option can be sold for a higher price.*

Scenario based risk measures.

In the scenario-based approach to measuring the risk of a portfolio one considers a number of possible future risk-factor changes (scenarios): such as a 10% rise in key exchange rates or a simultaneous 20% drop in major stock market induces or a simultaneous raise of key interest rates around the globe. The risk of the portfolio is then measured as the maximum loss of the portfolio under all scenarios, where certain extreme scenarios can be down-weighted to mitigate their effect on the result.

We now give a formal description. Fix a set $\mathcal{X} = \{x_1, \ldots, x_n\}$ of risk-factor changes (the scenarios) and a vector $\omega = (\omega_1, \ldots, \omega_n)^t \in [0, 1]^n$ of weights. Consider a portfolio of risky securities and denote by $l_{[t]}$ the corresponding loss operator. The risk of this portfolio is then measured as

$$\Phi_{[\mathcal{X},\omega]} := \max\{\omega_1 l_{[t]}(x_1), \ldots, \omega_n l_{[t]}(x_n)\}. \tag{2.17}$$

Many risk measures used in practice are of the form mentioned above.

Scenario based risk measures are a very useful risk-management tool for portfolios exposed to a relatively small set of risk factors. Moreover, they provide useful complementary information to measures based on statistics of the loss distribution. The main problem is of course to determine an appropriate set of scenarios and weighting factors. Moreover, comparison across portfolios which are affected by different risk factors is difficult.

2.4. Dataset of JPN 21

Our test-portfolio was created on the above mentioned concepts and has a strong Japanese focus. See figure 2.2.

Figure 2.2[6] describes the weekly portfolio returns for each stock and bond as part of the portfolio. It was constructed in a way to back-test the Tohoku earthquake in 2011. DEPFA Bank demanded an analysis of certain industry sectors in order to develop a hedging strategy to get rid of earthquake risks.

The Japanese Topix is one of the rare indexes that provides such an overview, as we can see in 2.3.

The next step is to think of meaningful relations. For example, construction related industries could be helpful after a massive earthquake, but Real Estate will probably suffer due to the news around the crippled Fukushimi Daiichi powerplant.

In Figure 2.4 we see the Topix and different industries around the Tohoku earthquake on 11-Mar-2011. We see a speculative move for Construction and other industries following the general trend.

Based on this first analysis, we expand the regarded time horizon and we were able to identify Rubber products as a possible candidate.

Changing the time horizon we see rubber products outperforming other industries in Figure 2.5. But why? Now we need to further analyse the Topix. We will find out which companies are actually causing the ups and down we seen in the carts. We need to calculate correlations in order to receive meaningful results and more importantly, analyse the companies.

[6] Charts created by: http://www.jpx.co.jp/english/markets/indices/histidx

It turns out that the Rubber industry within the Topix is mainly driven by Bridgestone, a well known tire-maker – who is also global leader in developing and producing seismic isolators, aside from other rubber products which are needed to rebuild Japan.

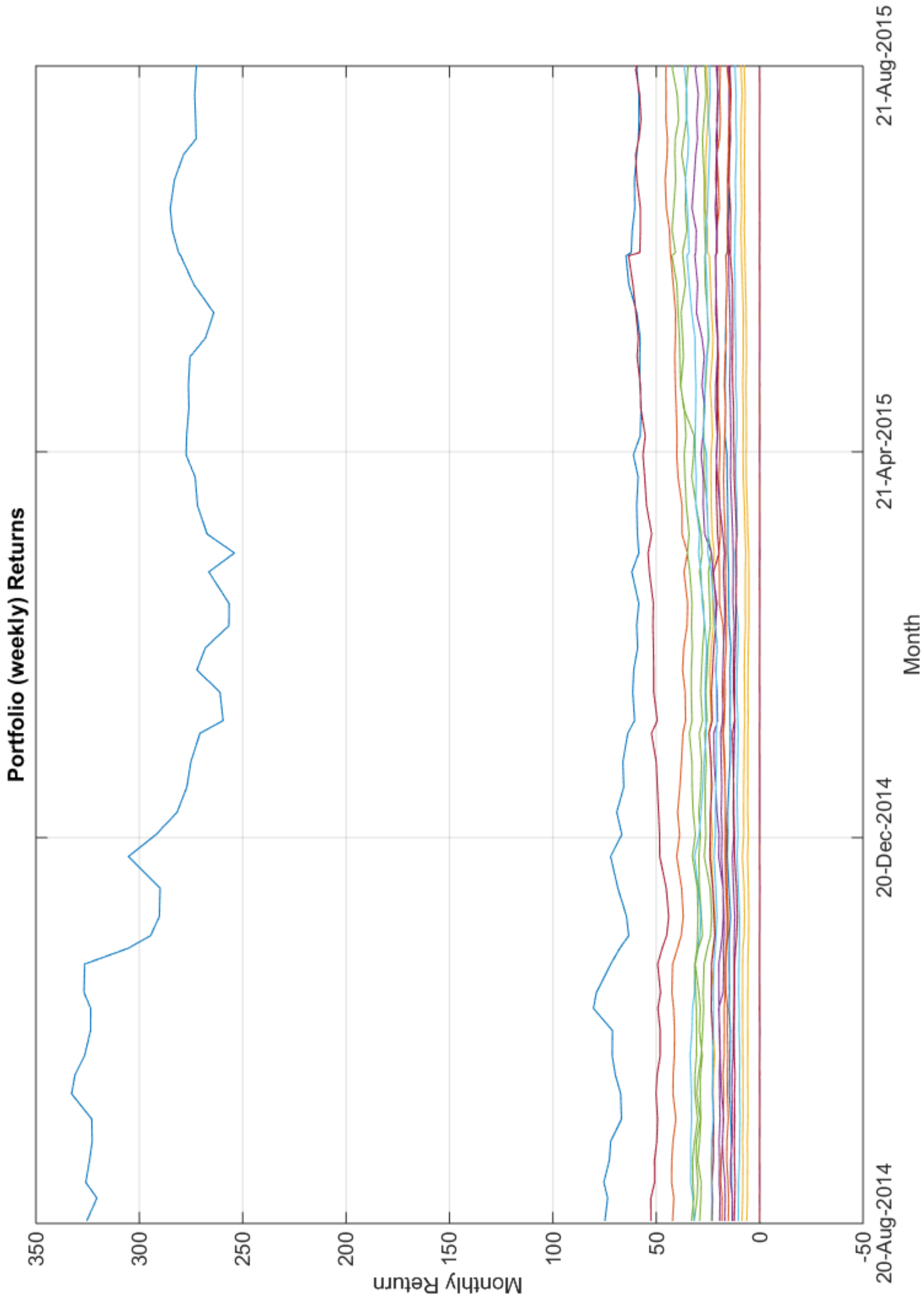

Figure 2.2.: Portfolio Returns

Industry	Manufacturing	Foods	Textiles & Apparels
		Pulp & Paper	Chemicals *
		Pharmaceutical *	Oil & Coal Products
		Rubber Products	Glass & Ceramics Products
		Iron & Steel	Nonferrous Metals
		Metal Products	Machinery
		Electric Appliances	Transportation Equipment
		Precision Instruments	Other Products
	Non-Manufacturing	Fishery, Agriculture & Forestry	Mining
		Construction	Electric Power & Gas
		Land Transportation	Marine Transportation
		Air Transportation	Warehousing & Harbor Transportation
			Services
		Information & Communication	Wholesale Trade *
		Retail Trade *	Banks *
		Securities & Commodities Futures *	Insurance *
		Other Financial Business *	Real Estate
		Services	

Figure 2.3.: Topix industries breakdown

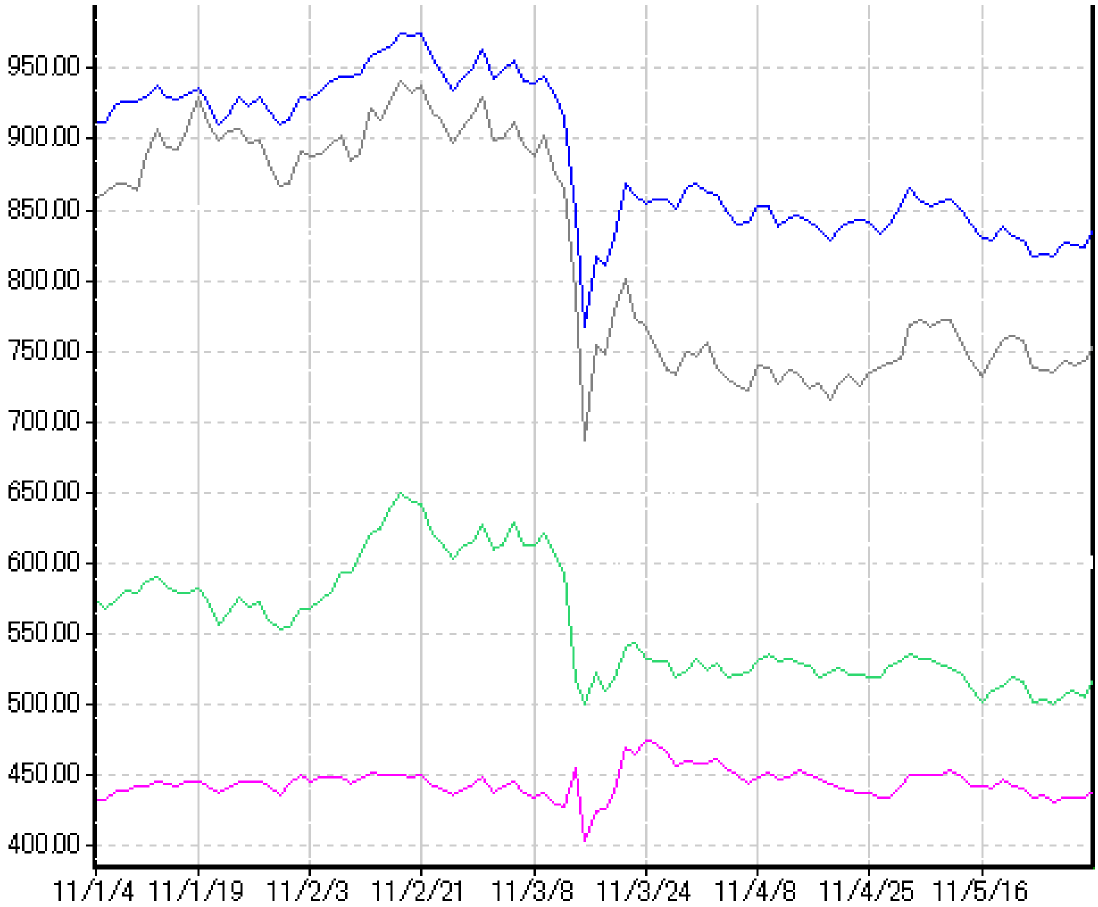

Figure 2.4.: Comparison of industries around Tohoku earthquake. From top to bottom: Topix, Real Estate, Insurance Companies, Construction

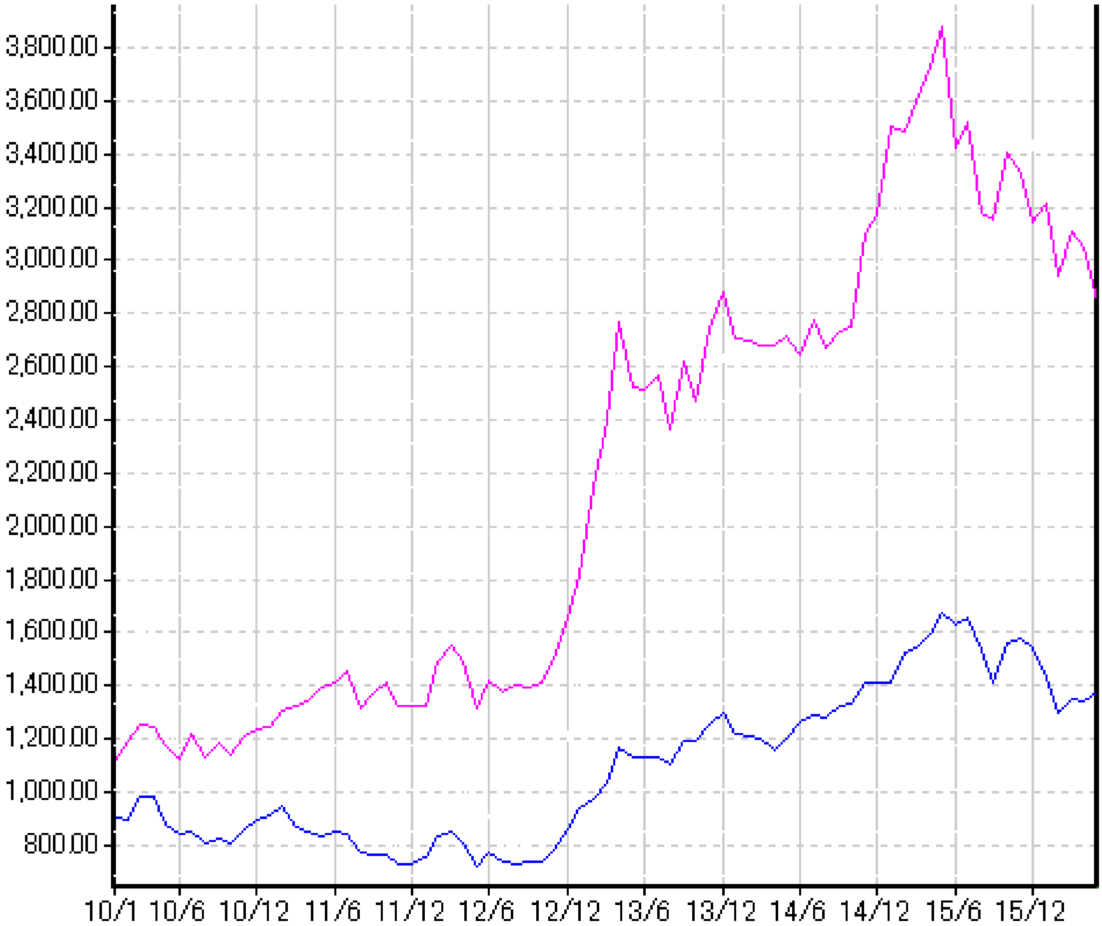

Figure 2.5.: Longterm view. From top to bottom: Rubber, Topix

Chapter 3.

Markowitz Theory

Before the 1950s, the desirability of an investment was mainly equated to its return. In his ground breaking publication in 1952, Harry Markowitz laid the foundation of the theory of portfolio selection by mapping the desirability of an investment onto a risk-return diagram, where risk was measured using standard deviation (See Markowitz [29] and [30]). Through the notion of an efficient frontier, the portfolio manager can optimize the return for a given risk level. We will also be following Atzberger [6] again with his introduction to the theory.

3.1. Efficient Frontier

The efficient frontier is the curve that shows all efficient portfolios in a risk-return framework.

Definition 16 (Efficient Portfolio)**.** *An efficient portfolio is defined as the portfolio, which maximizes the expected return for a given amount of risk (standard deviation), or the portfolio that minimizes the risk subject to a given expected return, (cf. 3.1).*

In Figure 3.1 we see the mapping of our risk versus our possible returns. All points on the line are optimal for their respective risk.

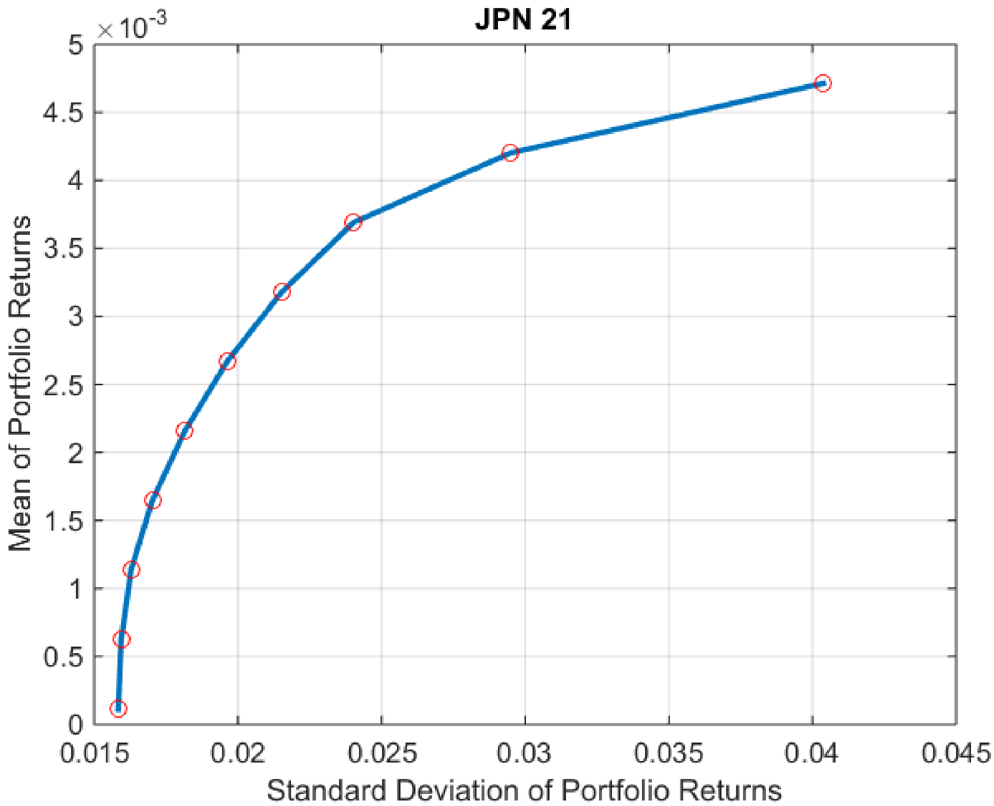

Figure 3.1.: Efficient Frontier - sometimes referred to as the Markowitz Bullet.

Markowitz Problem

With this understanding about the preferences of the investor, we shall consider a portfolio p to be desirable, if for a given expected rate of return μ_p the portfolio has the least variance σ_p^2. Finding such a portfolio is referred to as the Markowitz problem and can be stated mathematically as the following constrained optimization problem, recalling our earlier definitions of w_i as $w_i = \frac{W_i}{W}$ as the fraction of the wealth W_i invested in the ith asset:

Markowitz Problem

$$\min f(w_1, \ldots, w_n) = \frac{1}{2} \sum_{i,j=1}^{n} w_i w_j \sigma_{i,j} \tag{3.1}$$

To under stand this equation we need the following statement:

Proposition 2 (Markowitz Objective Function)**.** *The objective function is the variance of the portfolio*

Proof. The variance σ_p^2 for the rate of return of the portfolio is given by:

$$\sigma_p^2 = \mathbb{E}[|\rho_p - \mu_p|^2]$$

$$= \mathbb{E}\left[\left|\sum_{i=1}^{n} w_i(\rho_j - \mu_j)\right|^2\right]$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{n} w_i(\rho_j - \mu_j)\right)\left(\sum_{i=1}^{n} w_i(\rho_j - \mu_j)\right)\right]$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \mathbb{E}[(\rho_i - \mu_i)(\rho_j - \mu_j)]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{i,j} = \mathbf{w}^t V \mathbf{w}$$

□

The objective function is subject to:

$$g_1(w_1, \ldots, w_n) = \sum_{1=1}^{n} w_i\mu_i - \mu_p = 0$$

$$g_2(w_1, \ldots, w_n) = \sum_{1=1}^{n} w_i - 1 = 0$$

The first constraint specifies that the constructed portfolio is to have expected rate of return μ_p, while the second constraint arises from our definition of describing the investments in terms of relative values such as the rates of return of the assets and the relative portion of wealth invested in a given asset. To solve the constrained optimization problem, we shall use the method of Lagrange Multipliers. Many numerical methods also exist for these types of optimization problems. An introduction to such methods can be found in Tragler [45].

3.1.1. Optimal Portfolio of n Risky Assets

We now discuss the portfolios which are optimal in the sense that for a specified expected return the portfolio minimizes the variance of the return. To ensure that a well-defined solution exists, we shall make two assumptions about the collection of assets:

Assumption 1 (Assumption of a unique solution of the Markowitz problem)**.** *We need to keep two main assumptions in mind:*

- *The random returns are linearly independent, in the sense that any on return can not be expressed as a linear combination of the others,*
- *The expected return rates μ_i of the assets are not all equal.*

We shall now use the Method of Lagrange Multipliers to solve the Markowitz problem analytically. The Lagrangian is given by:

$$L(w, \lambda_1, \lambda_2) = \frac{1}{2}\mathbf{w}^t V \mathbf{w} + \lambda_1(\mu_p - \mathbf{w}^t \mu) + \lambda_2(1 - \mathbf{w}^t \mathbf{1}) \tag{3.2}$$

Remark 13 (Sufficient Conditions for Equally Constrained Problems)**.** *Assume the objective function f and constrains c_i are twice continuously differentiable in an open neighborhood of x_*, that $\nabla c_i(x_*)$ are linearly independent vectors at x_*, and that x_* is a local minimizer. Let $A(x_*) \in \mathbb{R}^{m\times n}$ be the Jacobian of all constraints at x_* (of full rank), and $Z(x_*) \in \mathbb{R}^{m\times(m-n)}$ be a basis matrix for the null space of $A(x_*)$. If there is a $\lambda^* \in \mathbb{R}^m$ such that the following conditions are satisfied at (x_*, λ^*), then x_* is a strict local minimizer:*

- $c_i(x_*) = 0$ *for* $i = 1, \dots, m$
- $\nabla_x L(x_*, \lambda^*) = 0 \Leftrightarrow Z(x_*)^t \nabla f(x_*) = \mathbf{0} \Leftrightarrow \nabla f(x_*) = A(x_*)^t \lambda_*$
- $Z(x_*)^t \nabla^2_{xx} L(x_*, \lambda^*) Z(x_*)$ *is positive definite*

To find the critical points of the Lagrangian, and hence the optimal portfolio, we must solve the first order equations.

$$\nabla_w L = V w_p - \lambda_1 \mu - \lambda_2 1 = 0 \tag{3.3}$$

and

$$\frac{\partial L}{\partial \lambda_1} = \mu_p - w_p^t \mu = 0$$
$$\frac{\partial L}{\partial \lambda_2} = 1 - w_p^t \mathbf{1} = 0$$

From 3.3 we conclude:

$$w_p = \lambda_1(V^{-1}\mu) + \lambda_2(V^{-1}\mathbf{1}),$$

where V denotes the covariance matrix.

From the first order equations, we get

$$(\mu^t V^{-1}\mu)\lambda_1 + (\mu^t V^{-1}\mathbf{1})\lambda_2 = \mu_p$$

$$(\mathbf{1}^t V^{-1}\mu)\lambda_1 + (\mathbf{1}^t V^{-1}\mathbf{1})\lambda_2 = 1.$$

Using these results, we can express these equations with

$$(\mu^t V^{-1}\mathbf{1}) = (\mu^t V^{-1}\mathbf{1})^t = \mathbf{1}^t (V^{-1})^t \mu = \mathbf{1}^t V^{-1}\mu$$

as

$$\begin{pmatrix} B & A \\ A & C \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \mu_p \\ 1 \end{pmatrix},$$

where

$$\begin{pmatrix} B & A \\ A & C \end{pmatrix} := \begin{pmatrix} \mu^t V^{-1}\mu & \mu^t V^{-1}\mathbf{1} \\ \mu^t V^{-1}\mathbf{1} & \mathbf{1}^t V^{-1}\mathbf{1} \end{pmatrix}. \tag{3.4}$$

At this point, we have to require a non-zero determinant to ensure a solution:

$$D = BC - A^2 \neq 0$$

This is indeed the case, if V is positive definite. In this case, also the inverse V^{-1} is positive definite and we can follow that $D \neq 0$ as a consequence of the positive definiteness of V.

Let us consider the vector

$$A\mu - B\mathbf{1} = \mu^t V^{-1}\mathbf{1}\mu - \mu^t V^{-1}\mu\mathbf{1},$$

which cannot vanish, since the only zero is $\mu = \mathbf{1}$, which is forbidden by our second assumption that μ_i are not all equal. Under this assumption, we can hence assume that $A\mu - B\mathbf{1} \neq 0$. From the positive definiteness of V we get V^{-1} positive definite and we know that:

$$
\begin{aligned}
0 &< (A\mu - B\mathbf{1})^t V^{-1}(A\mu - B\mathbf{1}) = \\
&= A^2\mu^t V^{-1}\mu - AB\mu^t V^{-1}\mathbf{1} - BA\mathbf{1}^t V^{-1}\mu + B^2\mathbf{1}^t V^{-1}\mathbf{1} \\
&= B^2C - A^2B = B(BC - A^2).
\end{aligned}
$$

From the fact that $B = \mu^t V^{-1}\mu > 0$, by positive definiteness of V we have that $D = BC - A^2 > 0$. In particular, $D \neq 0$, and we can invert the matrix in Equation 3.4 to obtain:

$$
\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \frac{1}{D}\begin{pmatrix} C & -A \\ -A & B \end{pmatrix}\begin{pmatrix} \mu_p \\ 1 \end{pmatrix}.
$$

More explicitly, this gives the solution for λ_1 and λ_2:

$$
\lambda_1 = \frac{C\mu_p - A}{D} \tag{3.5}
$$

$$
\lambda_2 = \frac{B - A\mu_p}{D} \tag{3.6}
$$

The weights of the portfolio can be found by substitution into our formerly introduced formula of w_p, which results in

$$
w_p = \left(\frac{C\mu_p - A}{D}\right)V^{-1}\mu + \left(\frac{B - A\mu_p}{D}\right)V^{-1}\mathbf{1} = \frac{1}{D}\left(BV^{-1}\mathbf{1} - AV^{-1}\mu\right) + \frac{1}{D}\left(CV^{-1}\mu - AV^{-1}\mathbf{1}\right)\mu_p. \tag{3.7}
$$

Notation 1. *For easier understanding of the formulas, we are introducing the following abbreviations:*

$$
g := \frac{1}{D}\left(BV^{-1}\mathbf{1} - AV^{-1}\mu\right)
$$

$$
h := \frac{1}{D}\left(CV^{-1}\mu - AV^{-1}\mathbf{1}\right)
$$

from which we can now express our weights by:

$$
w_p = g + \mu_p h. \tag{3.8}
$$

We shall refer to the portfolio which minimizes the variance for a specified expected rate of return μ_p as a frontier portfolio. We shall refer to the portfolio which has the minimum variance out of all frontier portfolios as the minimum variance portfolio and denote this by p^*. The solution shows explicitly how the the weights of the portfolio with minimum variance depend on the desired expected rate of return μ_p. In particular, it follows that all frontier portfolios can be attained by a linear combination of the two expressions g and h. In other words, instead of buying the individual assets, in theory one could obtain an equivalent rate of return of any frontier portfolio by investing in only two mutual funds (frontier portfolios) of the market.

Remark 14 (Covariance between random rates of frontier portfolios)**.** *Some interesting features of investing in portfolios can be derived from the theory. The covariance between the random rates of return ρ for two frontier portfolios ρ_a and ρ_b can be expressed in terms of μ_p^a and μ_p^b as:*

$$Cov[\rho_p^a, \rho_p^b] = \frac{C}{D}\left(\mu_p^a - \frac{A}{C}\right)\left(\mu_p^b - \frac{A}{C}\right) + \frac{1}{C} \tag{3.9}$$

3.1.2. Optimal Portfolios for n Risky Assets + 1 Risk-Free Asset

Extending the Markowitz problem by a risk-free asset, we follows Atzberger's approach [6]. In addition to the n risky assets, we now assume that there is one risk-free asset with return ρ_0. By risk-free we mean that the rate of return of the asset has zero variance[1]. This will make our model somewhat more realistic, since in actual markets an investor always has the opportunity to invest in an essentially risk-free treasury bond like a JGB or put money into a savings account.

Remark 15. *This is situation is can as well be described by the Mutual Fund Theorem. In portfolio theory, such a theorem stating that, under certain conditions, any investor's optimal portfolio can be constructed by holding each of certain mutual funds in appropriate ratios, where the number of mutual funds is smaller than the number of individual assets in the portfolio.*

In practice, however, things are more complicated, since this presupposes that investors only care about the expected returns, variances, and covariances of the risky assets.

[1]Discussions on what can be seen as risk-free assets are not part of this thesis.

In this case, the portion of wealth invested in the n risky assets no longer satisfies $\sum_{i=1}^{n} w_i = 1$, since we can always invest the remaining fraction of wealth in the risk-free asset or borrow funds. In the case that

$$\sum_{i=1}^{n} w_i < 1, \tag{3.10}$$

we say that we are under budget and invest the remaining portion of wealth in the risk-free asset. In the case of

$$\sum_{i=1}^{n} w_i > 1, \tag{3.11}$$

we say that we are over budget and borrow at the risk-free rate the excessive portion of wealth invested in the n risky assets.

This leads to a reformulation of what is meant by an optimal portfolio. In this case, an investor is no longer constrained to invest all of his or her wealth in the n risky assets. The objective then becomes to find the portfolio with the minimum variance for a specified expected rate of return. This results in:

$$\min f(w_1, \ldots, w_n) = \frac{1}{2} w^t V w \tag{3.12}$$

Subject only to the constraint such that the expected return is attained:

$$g(w_1, \ldots, w_n) = \rho_0 + w^t(\mu - \rho_0 \mathbf{1}) - \mu_{p^+} = 0 \tag{3.13}$$

with an efficient frontier portfolio and given expected return μ_{p^+}. We can apply the Langrange's method again and get the Langrangian:

$$L(w, \lambda) = \frac{1}{2} w^t V w - \lambda \left(\rho_0 - \mu_p + w^t(\mu - \rho_0 \mathbf{1})\right). \tag{3.14}$$

The condition that (w, λ) is a critical point transforms to:

$$\begin{aligned}
\nabla_w L &= V w_{p^+} - \lambda(\mu - \rho_0 \mathbf{1}) = 0 \\
\frac{\partial L}{\partial \lambda} &= \rho_0 - \mu_{p^+} + w_{p^+}^t(\mu - \rho_0 \mathbf{1}) = 0.
\end{aligned}$$

Using these results, we can easily express w_{p^+} and λ as follows:

$$w_{p^+} = \lambda V^{-1}(\mu - \rho_0 \mathbf{1})$$

$$\lambda = \frac{\mu_{p^+} - \rho_0}{(\mu - \rho_0 \mathbf{1})^t V^{-1}(\mu - \rho_0 \mathbf{1})}$$

By introducing

$$H := (\mu - \rho_0 \mathbf{1})^t V^{-1}(\mu - \rho_0 \mathbf{1})$$

we obtain in case $H > 0$:

$$w_{p^+} = \left(\frac{\mu_{p^+} - \rho_0}{H}\right) V^{-1}(\mu - \rho_0 \mathbf{1}).$$

Proof. We can now show, that in fact $H > 0$. Multiplying out the terms in the above equation, we receive

$$H = (\mathbf{1}^t V^{-1} \mathbf{1})\rho_0^2 - 2(\mathbf{1}^t V^{-1} \mu)\rho_0 - \mu^t V^{-1} \mu. \tag{3.15}$$

Remembering our definitions of A, B, C from the Markowitz Model, we obtain the expression

$$H = C\rho_0^2 - 2A\rho_0 + B, \tag{3.16}$$

which is obviously a quadratic equation in ρ_0. The value of H has the same sign for all values of ρ_0 if and only if there are no real roots. No real roots for a quadratic equation only occur, if the discriminative is negative:

$$(2A)^2 - 4CB = 4(A^2 - BC) < 0.$$

This can be shown to hold by the positive definiteness of V^{-1} and follows immediately from our earlier analysis of the inverse. □

To allocate our capital best, we can use the efficient frontier. It gives us a curve with optimal risk-return relations. We can use this information and compare it to our utility function. With these two functions, we can calculate an intersection to determine "our" best solution.

Example 7. *Let's take a look at the performance of this optimization, using historic Bloomberg data. Ahead of the big Tohoku earthquake and after it happened. The*

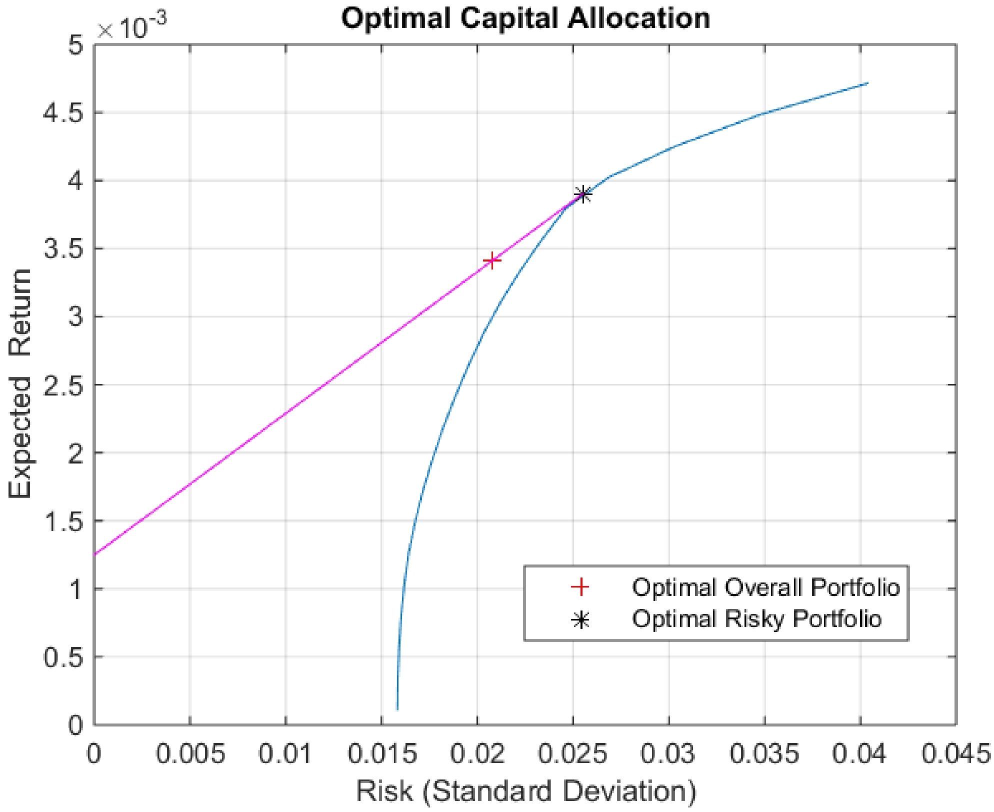

Figure 3.2.: Optimal Capital Allocation

following example provides a small and aggressive portfolio which is focused on the Japanese market with a very short, two-year time horizon. The portfolio consists of 10-year JGBs, Komatsu Inc. to represent construction-related companies, and Bridgestone which is listed as the biggest rubber company within the Topix industry groups and has a high correlation to cars as well. Japan Tobacco is used to diversify the risk. The Nikkei is used as a benchmark of the Japanese stock market.

From Table 3.3 we can see the stocks' last price and the 10-year JGB yield.

	JTobacco	**Bridgestone**	**Komatsu**	**10yr JGB**
Mean Return	0,04321	0,01979	0,05811	-0,03589

Table 3.1.: Mean Returns

	JTobacco	Bridgestone	Komatsu	10yr JGB
Q1 2010	1634	1447	1820	1,330
Q2 2010	1635	1574	1906	1,301
Q3 2010	1740	1596	1960	1,081
Q4 2010	1252	1443	1972	0,937
Q1 2011	1428	1476	1709	1,224
Q2 2011	1570	1770	2840	1,217
Q3 2011	1390	1547	1816	1,088
Q4 2011	1978	1905	1976	1,056
Q1 2012	1390	1542	1938	0,970
Q2 2012	2218	1905	2315	0,902
Q3 2012	1427	1542	2318	0,753
Q4 2012	2206	1859	1672	0,784

Table 3.2.: Quaterly Stock Prices

We can now calculate the quarterly returns and their average:

	JTobacco	Bridgestone	Komatsu	10yr JGB
Mean Return	0,04321	0,01979	0,05811	-0,03589

Table 3.3.: Mean Returns

Calculating the Covariance-Matrix and solving the non-linear equation, we receive a minimized portfolio variance of 3.16% with a desired return of 4% and we should allocate 12.5% into Japan Tobacco, 42.4% in Bridgestone, and 45.1% in Komatsu. Japanese Government Bonds are neglected as their return is too low. Comparing this allocation to the next quarter, we would have experienced a 4.7% increase, above our expectation and above the Nikkei, which increased 3.8% that quarter.

The example offers several insights to various problems. Although any value of return would be possible, using the historical calculated returns, only they are used, taking their variance into account. A 10% return can therefore not be calculated using this dataset. Benchmarking it on the next years data, we would make a loss of more than 1%, but it needs to be mentioned that the portfolio is not very diversified and investments should be seen on a long-term. Nevertheless, every portfolio manager will be measured on a yearly basis. We tackle this problem in a final outlook of the Fund Manager's Problem at the end of this work.

Chapter 4.

Merton's Portfolio Problem

Merton's portfolio problem is a well known problem in continuous-time finance and in particular inter-temporal portfolio choice. An investor must choose how much to consume and must allocate his wealth between stocks and a risk-free asset so as to maximize expected utility. We also advance by reducing our historical risks with this approach. The problem was formulated and solved by Robert C. Merton in [33] – both for finite lifetimes and for the infinite case. We follow his work in this chapter to a wide extent.

4.1. Problem Statement

The investor lives from time 0 to time T. Her/his wealth at time t is denoted by W_t. He starts with a known initial wealth W_0 (which may include the present value of wage income). At time t s/he must choose what amount of her/his wealth to consume, c_t, and what fraction of wealth to invest in a stock portfolio, π_t, while the remaining fraction $1 - \pi_t$ is invested in the risk-free asset.

The objective then is

$$\max_{c_s,\pi_s} \mathbb{E}\left[\int_0^T e^{-\rho s} u(c_s)\, ds + e^{-\rho T} u(W_T)\right] \tag{4.1}$$

with u denoting a utility-function[1] and ρ being discount rate.

[1] See Subsection 2.1.1 on risk aversion and utility functions.

The wealth evolves according to the stochastic differential equation (SDE)

$$dW_t = [(r + \pi_t(\mu - r))W_t - c_t]\,dt + W_t\pi_t\sigma\,dB_t \tag{4.2}$$

with r as the risk free rate, μ the expected return of the stock portfolio and σ the volatility of the stock market. A Brownian motion B_t is used to represent the stochastic term of this SDE. An introduction to this important stochastic process for continuous-time mathematical finance can be found in Schmock [43].

4.2. Dynamic Portfolio Choice

In this chapter, we follow Korn and Korn and their book on Modern Financial Mathematics [26] which gives a great introduction to this theory. In addition, Merton et al.'s work [34] optimum consumption and portfolio rules in a continuous-time model are used with Hamilton-Jacobi-Bellman equations, which are part of any introductory lecture on Applied Operations Research as presented by Tragler [46].

Let's begin this chapter with Bellman's principle of optimality:

"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."[2]

4.2.1. General Assumptions

We consider a market with constant market coefficients r, b,, σ. Let $\sigma \in \mathbb{R}^{d,m}$ with $d \leq m$ and denote σ as of full rank.

Remark 16. *We therefore do not require complete markets, meaning that we do not have to "fear" arbitrage. An arbitrage is a trading strategy that begins with no money, has zero probability of losing money, and has a positive probability of making money.*

[2] See Bellman [8], 1957, Ch. III.3

Basic Concept

Merton's basic idea is to describe the wealth function $X(t)$ at time t of an investor with a strategy (π, c) with a stochastic controlled differential equation

$$dX^u(t) = \mu(t, X^u(t), u(t), u(t))dt + \sigma(t, X^u(t), u(t))dW(t) \tag{4.3}$$

with μ, σ, u as follows:

$$u = (u_1, u_2) := (\pi, c)$$

$$\mu(t, x, u) = (r + u_1'(b - r \cdot \underline{1}))x - u_2$$

$$\sigma(t, x, u) = x u_1' \sigma$$

Hence, the choices of consumption and portfolio processes generate the controls. As there is a high number of verification theorems, we start with taking a deeper look of the standard basic problem.

4.2.2. Optimal Consumption and Optimal Terminal Wealth with Finite Time Horizon

Our goal is to maximize

$$J(t, x; u) := \mathbb{E}^{t,x}\left(\int_t^T U_1(t, u_2(t))dt + U_2(X^u(T)) \right)$$

by a choice of $u = (\pi, c)$. Hence,

$$V(t, x) = \sup_{u \in \mathbf{A}(t,x)} J(t, x; u) \tag{4.4}$$

describes the wealth or value of our portfolio as of time t with $\mathbf{A}(x)$ the set of all feasible controls for the utility function u.

The corresponding HJB equation can then be described as

$$\begin{aligned} \max_{u_1 \in [\alpha_1, \alpha_2]^d, u_2 \in [0,\infty)} & \frac{1}{2} u_1' \sigma\sigma' x^2 V_{xx}(t,x) \\ + & \left((r + u_1'(b - r \cdot \underline{1}))x - u_2 \right) V_x(t,x) \\ & + U_1(t, u_2) + V_t(t,x) = 0 \end{aligned} \tag{4.5}$$

with $V(T,x) = U_2(x)$ for given $-\infty < \alpha_1 \leq \alpha_2 < +\infty$.

Theorem 2 (Solution of the Portfolio Problem). *The portfolio problem*

$$\max_{(\pi,c) \in \mathbf{A}(x)} \mathbb{E} \left(\int_0^T e^{-\beta t} \frac{1}{\gamma} c(t)^\gamma dt + \frac{1}{\gamma} X(T)^\gamma \right) \tag{4.6}$$

is solved by strategy (π^*, c^*) *with*

$$\pi^* = \frac{1}{1-\gamma} (\sigma\sigma')^{-1} (b - r \cdot \underline{1}) \tag{4.7}$$

$$c^*(t) = X(t) \left(e^{\beta t} f(t) \right)^{\frac{1}{\gamma - 1}} \tag{4.8}$$

with $f(t)$ *given by*

$$\left(e^{\frac{a_1(T-t)}{1-\gamma}} + \frac{1-\gamma}{\gamma(a_1 - \beta)} \left(e^{\frac{(a_1-\beta)T}{1-\gamma}} - e^{\frac{(a_1-\beta)t}{1-\gamma}} \right) e^{-\frac{a_1 t}{1-\gamma}} \right)^{1-\gamma}.$$

Please find further details in Korn and Korn [26].

Chapter 5.

Aggregate Risk

This chapter is devoted to a number of theoretical concepts in quantitative risk management that fall under the broad heading of aggregate risk and mainly follows the work of McNeil et al. [32].

We understand aggregate risk as the risk of a portfolio, which could even be the entire position in a risky asset of a financial enterprise. The material builds on general ideas in risk measurement and also uses copula theory as well as some facts about elliptical distributions.

We treat the issue of general assumptions measuring aggregate risk. We discuss properties, that a good measure of risk should have with particular emphasis on aggregation properties and on industry standards. This leads us to study Value at Risk (VaR) and coherent risk measures. We as well consider the problem of bounding an aggregate risk, if we know something about the individual risks that contribute to the whole but have only limited information about their dependence. We discuss specific difficulties that arise when risk is measured with non-sub-additive risk measures like the VaR.

5.1. Coherent Measures of Risk

Using economic reasoning, Artzner et al. [4] proposed a list of properties that a good risk measure should have. They specified a number of axioms that any so-called coherent risk measures should satisfy. Moreover, they studied the coherence properties of widely used risk measures like the VaR or expected shortfall and gave a characterization of all coherent risk measures in terms of generalized scenarios. This work will therefore follow their approach.

The Axioms of Coherence

In order to introduce the axioms of coherence we have to give a formal definition of risk measures. Fix some probability space (Ω, F, P) and a time horizon Δ. We denote by $L^0(\Omega, F, P)$ the set of all r isk values on (Ω, F), which are almost surely finite. Financial risks are represented by a set $M \subset L^0(\Omega, F, P)$ of risk values, which we interpret as portfolio losses over some time horizon Δ. The time horizon is left unspecified and will only enter when specific problems are considered. We often assume that M is a convex cone, i.e., that $L_1 \in M$ and $L_2 \in M$ imply that $L_1 + L_2 \in M$.

Definition 17 (Convex)**.** *$C \in \mathbb{R}^n$ is convex, if*

$$\forall x, y \in C \quad \forall \lambda \in [0, 1] : \lambda x + (1 - \lambda)y \in C.$$

Definition 18 (Cone)**.** *$S \in \mathbb{R}^n$ with $0 \in S$. S is a cone, if*

$$x \in S \Rightarrow \forall \alpha > 0 : \alpha x \in S.$$

Obviously, if additionally S is convex, we call it a convex cone.

Risk measures are real-valued functions $\phi : M \rightarrow R$ defined on such cones of risk values, satisfying certain properties.

We interpret $\phi(L)$ as the amount of capital that should be added to a position with loss given by L, so that the position becomes acceptable to an external or internal risk controller. Positions with $\phi(L) \leq 0$ are acceptable without injection of capital; if $\phi(L) < 0$, capital may even be withdrawn. Note that our interpretation of L differs from that in [4], where risk values $L \in M$ is interpreted as the future value (instead of the loss) of a position currently held. This leads to some sign changes in the discussion of the axioms of coherence compared with other presentations in the literature. Also note that in order to simplify the presentation we set interest rates equal to zero so that there is no discounting.

Now we can introduce the axioms that a risk measure $\phi : M \rightarrow R$ on a convex cone M should satisfy in order to be called coherent.

Axiom 1 (Translation Invariance)**.** *For all $L \in M$ and every $l \in R$ we have $\phi(L+l) = \phi(L) + l$.*

Axiom 1 states that by adding or subtracting a deterministic quantity l to a position leading to the loss L we alter our capital requirements by exactly that amount. The axiom is in fact necessary for the risk-capital interpretation of ϕ to make sense.

Remark 17. *Consider a position with loss L and $\phi(L) > 0$. Adding the amount of capital $\phi(L)$ to the position leads to the adjusted loss $\tilde{L} = L - \phi(L)$, with $\phi(\tilde{L}) = \phi(L) - \phi(L) = 0$, so that the position $\tilde{L}$ is acceptable without further injection of capital.*

Axiom 2 (Sub-additivity)**.** *For all $L_1, L_2 \in M$ we have $\phi(L_1 + L_2) \leq \phi(L_1) + \phi(L_2)$.*

The rationale behind this axiom is summarized by Artzner et al. [4] in the statement that "a merger does not create extra risk" (ignoring of course any problematic practical aspects of a merger!). It is also the most debated of the four axioms characterizing coherent risk measures, probably because it rules out VaR as a risk measure in certain situations. We provide some arguments explaining why sub-additivity is indeed a reasonable requirement.

Sub-additivity reflects the idea that risk can be reduced by diversification, a time honored-principle in finance and economics. In particular, we will see, that the use of non-sub-additive risk measure in a Markowitz-type portfolio optimization problem may lead to optimal portfolios that are very concentrated, which would be deemed quite risky by normal economic standards.

If a regulator uses a non-sub-additive risk measure in determining the regulatory capital for a financial institution, that institution has an incentive to legally break up into various subsidiaries in order to reduce its regulatory capital requirements – as it happened at DEPFA Bank. Similarly, if the risk measure used by on organized exchange in determining the margin requirements of investors in non-sub-additive, an investor could reduce the margins s/he has to pay by opening a different account for ever position in his portfolio.

Sub-additivity makes decentralization of risk-management systems possible.

Example 8. *Consider two trading desks with positions leading to losses L_1 and L_2. Imagine that a risk manager wants to ensure $\phi(L)$, the overall risk $L = L_1 + L_2$, is* *smaller than some number M. If s/he uses a risk measure ϕ, which is sub-additive, s/he may simply choose bounds M_1 and M_2 such that $M_1 + M_2 \le M$ and impose on each of the desks the constraint that $\phi(L_i) \le M_i$: sub-additivity of ϕ then ensures automatically that $\phi(L) \le M_1 + M_2 \le M$.*

Axiom 3 (Monotonicity)**.** *For $L_1, L_2 \in \mathbf{M}$ such that $L_1 \le L_2$ almost surely we have $\phi(L_1) \le \phi(L_2)$.*

From an economic viewpoint this axiom is obvious: positions that lead to higher losses in every state of the world require more risk capital.

Axiom 4 (Convexity)**.**

$$\phi(\lambda L_1 + (1-\lambda)L_2) \le \lambda\phi(L_1) + (1-\lambda)\phi(L_2)$$

with $\lambda \in [0,1]$ for all $L_1 L_2 \in \mathcal{M}$.

For a risk measure satisfying Axioms 1 and 2, the monotonicity axiom is equivalent to the requirement that $\phi(L) \le 0$ for all $L \le 0$. To see this, observe that this axiom implied, that if $L \le 0$, then $\phi(L) \le \phi(0) = 0$. The latter equality follows by the monotonicity since $\phi(0) = \phi(\lambda 0) = \lambda\phi(0)$ for all $\lambda > 0$. Conversely, if $L_1 \le L_2$ and we assume that $\phi(L_1 - L_2) \le 0$, then $\phi(L_1) = \phi(L_1 - L_2 + L_2) \le \phi(L_1 - L_2) + \phi(L_2)$ by Axiom 2, which implies $\phi(L_1) \le \phi(L_2)$.

Definition 19. *A risk measure ϕ whose domain includes the convex cone $\mathcal{M}$ is called coherent (on $\mathcal{M}$) if it satisfied Axioms 1-4.*

Note that the domain is an integral part of the definition of a coherent risk measure. We will often encounter functionals on $L^0(\Omega, \mathcal{F}, P)$, which are coherent only if restricted to a sufficiently small convex cone $\mathcal{M}$.

Remark 18 (Convex Measures of Risk)**.** *Positive homogeneity has been criticized and, in particular, it has been suggested that for large values of the multiplied λ we should have $\phi(\lambda L) > \lambda\phi(L)$ to penalize a concentration of risk and the ensuing liquidity problems. As shown before, this is impossible for a sub-additive risk measure. This problem has led to the study of convex risk measures. In this class, the conditions of sub-additivity and positive homogeneity have been relaxed; instead one requires only the weaker property of convexity.*

The economic justification of this axiom is again the idea that diversification reduces risk. Within the class of convex risk measures it is possible to find risk measures penalizing concentration of risk in the sense that $\phi(\lambda L) \geq \phi(L)$ for $\lambda > 1$. Convex risk measures have recently attracted a lot of attention.

In Artzner et al. [4], we face the problem of defining a complete set of axioms that have to be fulfilled by a measure of risks in a generalized sense. A measure which satisfies these axioms is defined as *Coherent Measure of Risk*. It is then shown that whenever a portfolio is undoubtedly riskier than another one, it will always have a higher value of risk if the measure is coherent. On the contrary, any measure which does not satisfy some of the axioms will produce paradoxical results of some kind giving a wrong assessment of relative risks. In the paper, the class of coherent measures is identified and characterized and a coherence test is given. In the second axiom for a measure, we find a familiar concept:

While Artzner et al. first only considered discrete probability spaces, Artzner et al. [5] later extended their definitions to the case of arbitrary probability spaces. A coherent risk measure $\rho(X)$ of an investment with the random return X is a real-valued function defined on the space of real-valued random variables which satisfies the following axioms:

1. Translation invariance
2. Sub-additivity
3. Positive homogeneity
4. Positivity

In short, think of ρ as representing some risk measure, and $\rho(x)$ as the risk of asset x under that measure. If sub-additivity holds, then the risk of holding assets 1 and 2 simultaneously must be less than or equal to the sum of their individual risks: $\rho(x_1 + x_2) \leq \rho(x_1) + \rho(x_2)$.

For example, volatility (standard deviation) is a sub-additive risk measure. We know this intuitively from diversification: a portfolio is less volatile than the sum of its component volatilities.

As it relates to finance, sub-additivity is one of the four axioms characterizing coherent measures of risk. This class of risk measures was introduced in Artzner et al. [4]. Think of these as risk measures with desirable properties that will not be subverted by strange-behaving portfolios. It's important to note that sub-additivity is not a statement of fact – it is easy to define risk measures that are not sub-additive – but rather an axiom that risk measures must satisfy in order to be coherent.

Artzner et al. described additivity nicely as the idea that "a merger does not create extra risk," and lists a number of practical points which follow from it. An interesting one is that if risk were not sub-additive, then a person wanting exposure to asset 1 and asset 2 would be better off opening a separate account for each asset, as the (risk-based) margin requirement would be lower than if s/he held both in the same account.

Remark 19. *A convex risk measure is coherent if it satisfies positive homogeneity.*

The most (in)famous risk measure that does not satisfy this axiom is VaR. The VaR of a portfolio of two assets can be greater than the sum of their individual VaRs. This is because VaR is a quantile-based measure as further discussed in Artzner et al. [4].

5.2. Value at Risk

The following chapter is mainly based on the work of Gaivoronski and Pflug [19] and on McNeil et al. [32], as well Lectures Notes on CreditRisk+ by Schmock [42].

Value-at-Risk is probably the most widely used risk measure in financial institutions and also got accepted in the Basel II and Basel III capital-adequacy framework.

Definition 20 (Value-at-Risk)**.** *Given some confidence level $\alpha \in (0,1)$, the VaR of our portfolio at the confidence level α is given by the smallest number l such that the probability (described by distribution F) that the loss L exceeds l is no larger than $(1-\alpha)$. Formally:*

$$VaR_\alpha = \inf\{l \in \mathbb{R} : P[L > l] \leq 1-\alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\} \tag{5.1}$$

Therefore, the VaR can be seen as a quantile of the loss distribution with typical values $\alpha = 0.95$ or $\alpha = 0.99$. A typical time horizon in market risk is about 10 days and in credit risk one year. Figure 5.1 shows the loss distribution of our JPN 21 portfolio. We will now use it for our further calculations.

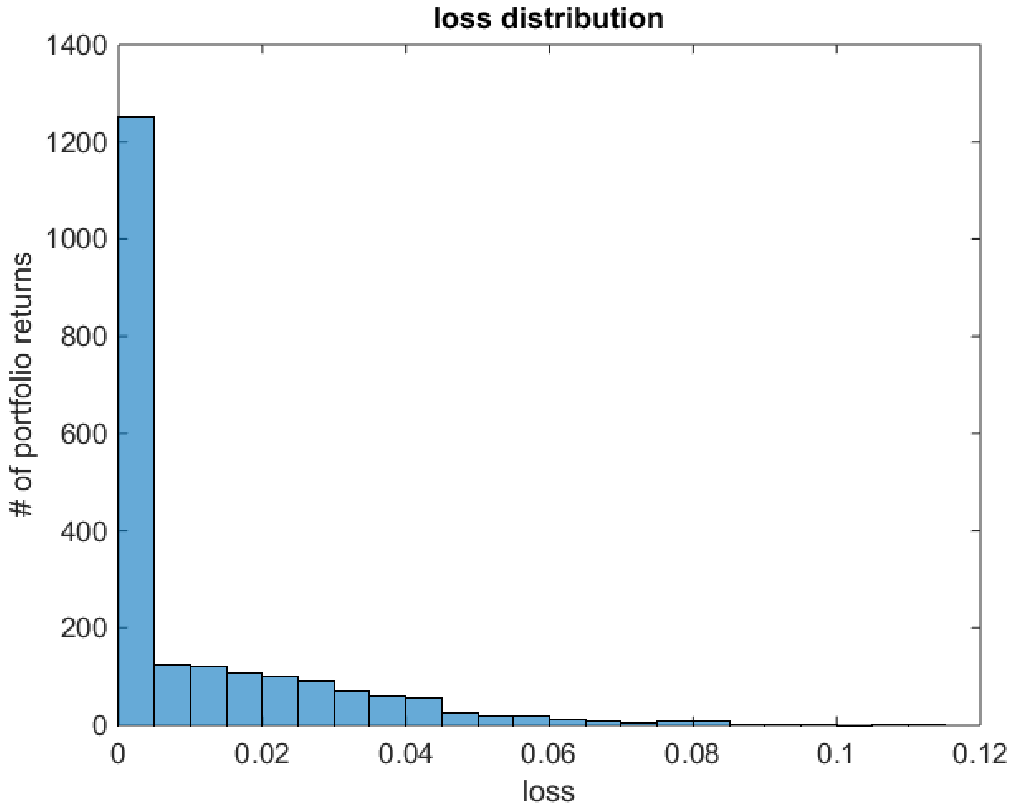

Figure 5.1.: Loss Distribution

It is easily seen that the VaR is translation invariant, positive homogeneous and monotone on $L^0(\Omega, \mathcal{F}, P)$.

Proposition 3 (VaR Properties). *Value at Risk is translation invariant, positive homogeneous and monotone.*

However, as the following discussions proves, the sub-additivity property fails to hold for VaR in general and therefore, VaR is not a coherent risk measure in general. The following example illustrates that measuring risk with VaR can lead to nonsensical results.

Definition 21 (Default-able Bonds and Loans). *A defaultable (corporate) bond is a pure discount bond which promises to pay one yen at its fixed expiry date. Since we imagine being a risk manager and promises are not always kept, we define a default-able bond as a fictitious financial security which pays one yen at a fixed time T if no default occurs prior to that time and it pays a random recovery rate $R \in [0, 1)$ at default time t in $[t, T]$ otherwise.*

Loans on the other hand tend to be agreements between banks and customers. Loans are usually non trade-able and the bank is obliged to see out the term of the loan.

Remark 20. *The main difference between a bond and loan is that a bond is highly trade-able. If we buy a bond, there is usually a market where we can trade bonds. This means we can sell the bond, rather than wait to the end of a, let's say 30 year period. In practice, one would buy bonds when one wishes to increase one's portfolio in that way.*

Following Schmock's [42] approach, we see that contrary to its widespread use, VaR is not perfectly suitable as a risk measure for two economic reasons. First of all, it does not take into account the size of losses, which occur with probability at most $1 - \delta$ for $\delta \in (0, 1)$, meaning that it disregards risks with high effects but low probability. Secondly, VaR is not sub-additive in general, i.e., it can happen that $VaR(X) + VaR(Y) < VaR(X + Y)$ for loss variables X and Y, meaning that diversification might seem to increase risk when it is measured with VaR, see the following example.

Example 9 (VaR is not sub-additive)**.** *Consider a loan of 100 JPY with default probability $p = 0.8\%$, which leads to a VaR at level 1% of zero. On the other hand, if we consider two independent loans of 50 JPY each with the same default probability $p = 0.8\%$, then the probability of at least one default is $2p - p^2 > 0.96\%$ and thus the VaR at level 1% equals 50 JPY. This means we would prefer the 100 JPY loan as the safer investment, which contradicts the idea of diversification.*

According to Danielsson et al. [16], this problem is caused by the fact that VaR is a quantile on the distribution of profit and loss and not an expectation, so that the shape of the tail before and after the VaR probability need not have any bearing on the actual VaR number.

Unrecognized violations of VaR sub-additivity can have serious consequences for risk models. First, they can provide a false sense of security, so that a financial institution may not be adequately hedged. Second, it can lead a financial institution to make a suboptimal investment choice, if VaR, or a change in VaR, is used for identifying the risk in alternative investment choices.

In fact, for any coherent risk measure ϕ, which depends only on the distribution of L, we get due to the monotonicity

$$\phi\left(\sum_{i=1}^{100} L_i\right) \leq \sum_{i=1}^{100} \phi(L_i) = 100\phi(L_0) = \phi(100L_1).$$

Hence, any coherent risk measure, which depends only on the loss distribution, will lead to a higher risk capital requirement for let's say a portfolio A than for some portfolio B.

VaR is, however, sub-additive in the idealized situation where all portfolios can be represented as linear combinations of the same set of underlying elliptically distributed risk factors. In this case, both the marginal loss distribution of the risk factors and the copula possess strong symmetry.

Remark 21. *In McNeil et al.* [32]*, we learn that an elliptical model may be a reasonable approximate model for various kinds of risk-factor data, such as a stock or a exchange-rate returns.*

However, there are plenty of other methods similar to the VaR which address this problem.

Definition 22 (Mean-VaR)**.** *Be μ the mean of the loss distribution. For capital-adequacy purposes, the mean-VaR*

$$VaR_\alpha^{mean} := VaR_\alpha - \mu \tag{5.2}$$

is used. For small time-horizons, the mean-VaR is of little importance, whereas in loan portfolios to buffer unexpected losses, the mean-VaR is of high importance.

VaR optimization can be considered as a stochastic programming (SP) problem of special type. It is related to SP problems with probability constraints, as shown by Prekopa [40].

Suppose R is some risk measure like the VaR or the standard deviation. For a given minimal expected return μ, let us consider the following problem.

T

5.2.1. VaR

It has long been recognized, however, that there are several conceptual difficulties with using standard deviation as a measure of risk, as it is used in VaR calculations. To sum up the most important issues:

1. VaR could violate second order stochastic dominance and therefore does not always describe risk aversion in the traditional sense.
2. VaR is not smooth: Events with a probability just below 1% are not taken into account. This changes immediately if the probability is exactly or greater than 1%.
3. As mentioned above, VaR is not always sub-additive: If VaR is calculated for each unit within a bank, the sum of the Values at Risk of each unit could be lower(!) than the VaR for the whole bank.

   Obviously, this contradicts the idea of diversification, because risk could than be reduced by running each unit separately. (These centrifugal forces are presumably not in the interest of the top management.) The lack of sub-additivity makes VaR a problematic criterion for portfolio optimization, the internal allocation of capital.

It is therefore highly recommended to make decisions based not solely on the Value-at-Risk, but to take other risk measures into account.

5.2.2. Shortfall as a Risk Measure

This sections is mainly based on Bertsimas et al. [`?bertsimas2004`] and Schmock [42].

Motivated from second-order stochastic dominance, we introduce a risk measure that we call shortfall. We examine shortfall properties and discuss their relation to such commonly used risk measures as standard deviation, VaR, lower partial moments, and coherent risk measures. We show that the mean-shortfall optimization problem, unlike mean-VaR, can be solved efficiently as a convex optimization problem, while the sample mean-shortfall portfolio optimization problem can be solved very efficiently as a linear optimization problem.

Bertsimas et al. [`?bertsimas2004`] provided empirical evidence

a) in asset allocation, and

b) in a problem of tracking an index using only a limited number of assets that the mean-shortfall approach might have advantages over mean-variance.

5.3. Expected Shortfall as a Coherent Risk Measure

Definition 23 (Expected Shortfall; cf. [42])**.** *Let X be a real-valued random variable. Then the expected shortfall of the loss variable X at level $\delta \in (0,1)$ is defined as*

$$ES_{\delta}[X] = \frac{\mathbb{E}\left[X 1_{\{X > q_\delta(X)\}}\right] + q_\delta(X)(\mathbb{P}[X \leq q_\delta(X)] - \delta)}{1-\delta}.$$

using a quantile function ρ_δ for X.

Remark 22. *When expected shortfall is taken as a risk measure, then (contrary to VaR) the sizes of large losses exceeding the threshold $q_\delta(X)$ are clearly taken into account by this conditional average. The additional term in the definition is necessary to prove the sub-additivity of expected shortfall. The second representation justifies the name conditional value-at-risk (CVaR), which is used in the pertinent literature and will be used in this work when applicable.*

With our new tools, we can compare efficient frontiers with 5.2. We see very similar results for the CVaR and the MAD calculations, whereas the mean-variance approach underperforms. Not only does it undervalue our risk, it also gives us less opportunity to make profits.

Application in Credit Risk

In credit risk we model a portfolio of loans and calculate the overall risk, in terms of defaults and its resulting losses. Due to the enormous number of calculations as seen in basic Bernoulli models, poison approximation models are widely used as trade-off effects justify its use.

If we have an estimate for the Wasserstein distance of two distributions, see [42], then we get bounds for the expected shortfall of these distributions.

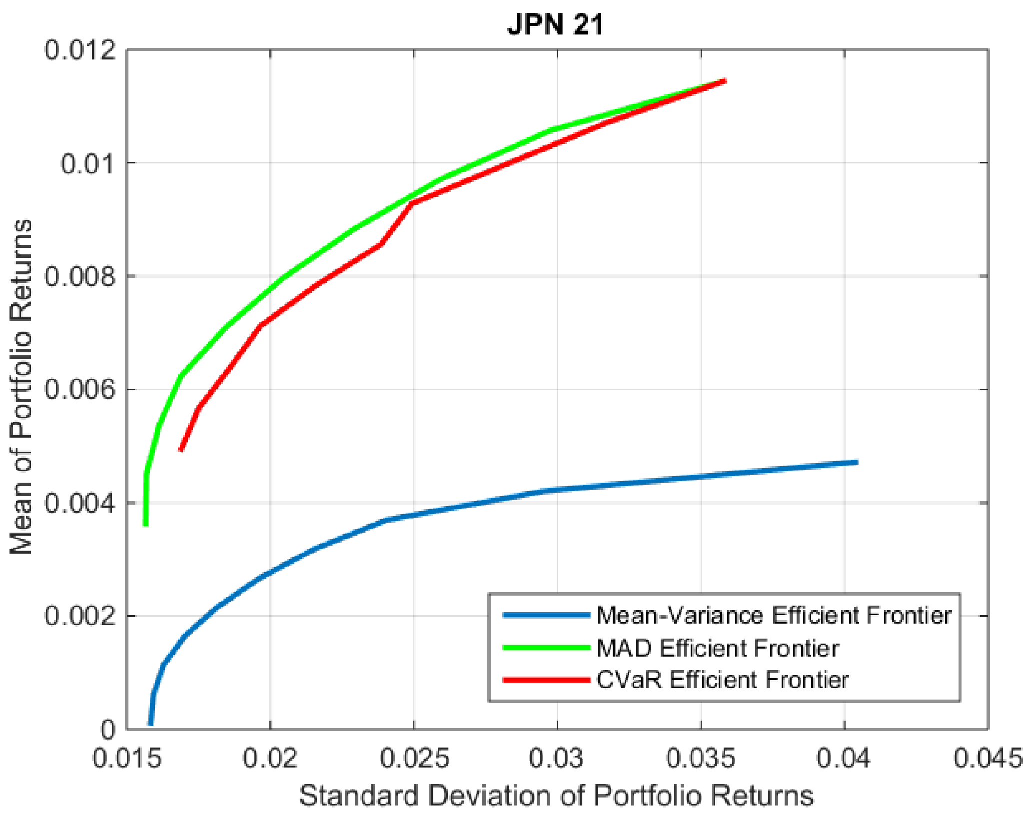

Figure 5.2.: Efficient Frontier Comparison

Lemma 2 (Expected Shortfall and Wasserstein Distance)**.** *Let X and Y be real-valued, integrable random variables and denote the Wasserstein distance of their distributions by $d_W(L(X), L(Y))$. Then the expected shortfall of X and Y satisfies, for every $\delta \in (0,1)$,*

$$|ES_\delta[X] - ES_\delta[Y]| \leq \frac{d_W(L(X), L(Y))}{1-\delta}.$$

Proof. A detailed proof can be found in [42]. □

Allocation of risk capital by the expected shortfall principle has a number of good properties. For an axiomatic approach to risk capital allocation, see Kalkbrener [24]. In contrast to our VaR-example, we can see that expected shortfall is definitely an improvement.

Remark 23 (Diversification and Expected Shortfall)**.** *Property $ES_\delta[X, Z] \leq ES_\delta[X, X]$ shows that X considered as a sub-portfolio of any other portfolio Z does not need more risk capital than on its own, meaning that diversification never increases the risk capital.*

5.4. Convex Risk Measures for Portfolio Optimization and Concepts of Flexibility

In this section, we follow the approaches of Luethi and Doege [?luthi].

Due to their axiomatic foundation and their favorable computational properties convex risk measures are becoming a powerful tool in financial risk management. We will review the fundamental structural concepts of convex risk measures within the framework of convex analysis. Then we will exploit them for deriving strong duality relations in a generic portfolio optimization context. In particular, the duality relationship can be used for designing new, efficient approximation algorithms based on Nesterov's smoothing techniques [37] for non-smooth convex optimization. Furthermore, the presented concepts enable us to formalize the notion of flexibility as the (marginal) risk absorption capacity of a technology or (available) resources.

Conditional Value-at-Risk

The most popular and widely used coherent risk measure is the Conditional Value-at-Risk (CVaR) as discussed in Rockafellar and Uryasev [41].

Definition 24 (Conditional Value-at-Risk (CVaR))**.** *Given the fact that Conditional Value-at-Risk is defined as the expected tail loss and based on our VaR definiton one would intuitively define CVaR with confidence level* $0 < \beta < 1$ *as*

$$CVaR_\beta(X) = E_{\mathbb{P}}[X|X > VaR_\beta(X)] \tag{5.3}$$

This popular definition of CVaR will in general result in a coherent measure if X has continuous distribution.

Figure 5.3 shows us the relative distribution of our portfolio weights. We experience trends as we focus on the best performers, depending on our constrains. All three ways however decide on the same stock to focus on. Mean-Variance and MAD seem to be more robust and similar.

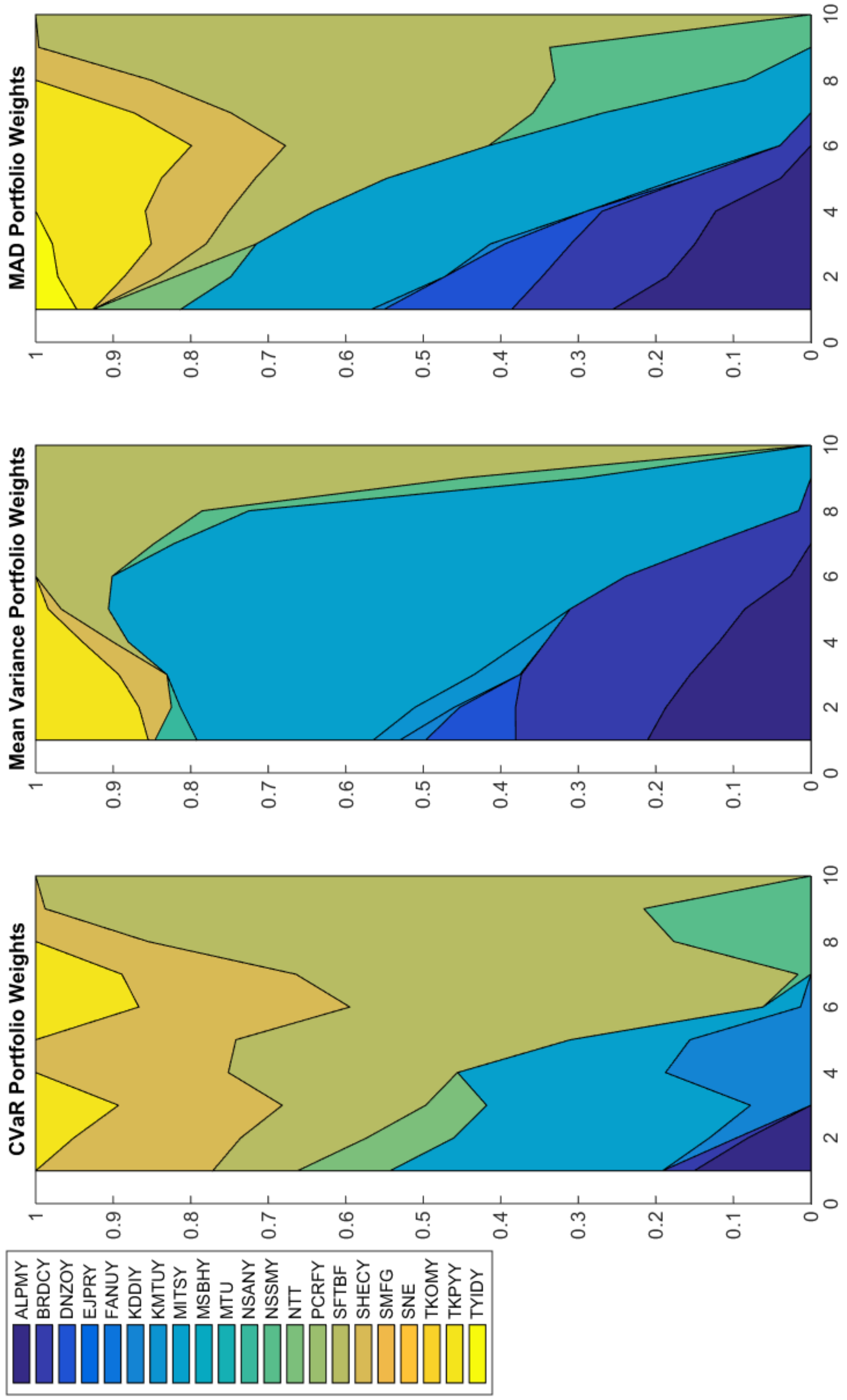

Figure 5.3.: Portfolio Weights Comparison

Chapter 6.

Optimal Control of Conditional Value-at-Risk

We now consider continuous-time stochastic optimal control problems featuring CVaR in the objective and follow Miller and Yang's approach as in [36]. The major difficulty in these problems arises from time-inconsistency, which prevents us from directly using dynamic programming. To resolve this challenge, we convert to an equivalent bi-level optimization problem in which the inner optimization problem is standard stochastic control. Furthermore, we provide conditions under which the outer objective function is convex and differentiable. We compute the outer objective's value via a Hamilton-Jacobi-Bellman equation and its gradient via the viscosity solution of a linear parabolic equation, which allows us to perform gradient descent. The significance of this result is that we provide an efficient dynamic programming-based algorithm for optimal control of CVaR without lifting the state-space. To broaden the applicability of the proposed algorithm, we provide convergent approximation schemes in cases where our key assumptions do not hold and characterize relevant sub-optimality bounds. In addition, we extend our method to a more general class of risk metrics, which includes mean-variance and median-deviation. We also demonstrate a concrete application to portfolio optimization under CVaR constraints. Our results contribute an efficient framework for solving time-inconsistent CVaR-based dynamic optimization.

CVaR takes into account the possibility of tail events where losses exceed VaR. Due to the superior mathematical properties and practical implications, CVaR has gained popularity in risk management. In particular, static optimization with CVaR functions can be efficiently performed via convex and linear programming methods. With the advances in optimization algorithms for CVaR, this risk measure has shown to be useful in various finance and engineering applications.

Dynamic or sequential optimization of CVaR is often of interest when decisions can be made at multiple stages. In such an optimal control setting, we can optimize a control action at a certain time based on the information from observations up to that time. This dynamic control approach enjoys an effective usage of information gathered in the process of making decisions under uncertainty. The need for efficient optimal control tools with CVaR is also motivated by emerging dynamic risk management methods in engineering and finance.

The major challenge in optimal control involving CVaR arises from its time-inconsistency. For example, an optimal strategy for tomorrow constructed today is no longer optimal when considered tomorrow because CVaR is not a time-consistent risk measure. Mathematically, this time-inconsistency prevents us from directly applying dynamic programming, in contrast with problems involving Markov risk measures or risk-sensitive criteria.

Following the research by Miller and Yang [36], we propose a method to solve continuous-time and continuous-space optimal control problems involving CVaR. By using a so-called extremal representation of CVaR, we reformulate the optimal control problem as a bilevel optimization problem in which the outer optimization problem is convex and the inner optimization problem is standard stochastic optimal control.

6.1. Problem Setup

6.1.1. Controlled Process

Let $(\Omega, F, \mathbb{P})$ be a probability space supporting a standard d-dimensional Brownian Motion W with associated filtration $\{F_t\}_{0 \le t \le T}$ satisfying the usual conditions. Let $\mathbb{A}$ be a compact and finite-dimensional set of controls. Define a set of $\mathbf{A}$ admissible strategies as the collection of all F_t-progressively measurable processes which are valued in $\mathbb{A}$ almost surely. The control $A \in \mathbf{A}$ affects a system state of interest through the following stochastic differential equation:

$$dX_t^A = \mu(X_t^A, A_t)dt + \sigma(X_t^A, A_t)dW_t \tag{6.1}$$
$$X_0^A = x_0 \in \mathbb{R}^n.$$

We assume that $\mu : \mathbb{R}^n \times \mathbb{A} \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \times \mathbb{A} \to \mathbb{R}^{n\times d}$ are continuous functions such that, for some $K > 0$,

$$\|\mu(x,a) - \mu(x',a)\| + \|\sigma(x,a) - \sigma(x',a)\| \le K\|x - x'\|$$

$$\|\mu(x,a)\| + \|\sigma(x,a)\| \le K(1 + \|x\| + \|a\|)$$

for all $x, x' \in \mathbb{R}^n$ and for all $a \in \mathbb{A}$. Under these conditions, for each control $A \in \mathbf{A}$, there exists a unique strong solution X^A, of the SDE.

6.1.2. Optimal Control with a Class of Risk Measures

The main goal of Miller and Yang's work was to provide an efficient algorithm for solving the following stochastic optimal control problem with a non-standard objective:

$$\inf_{A\in\mathbf{A}} \rho(g(X_T^A)) \tag{6.2}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a given convex cost function and $\rho : L^2(\Omega) \to \mathbb{R}$ is a class of risk measures as defined below.

Remark 24. *The formulation of the objective function includes the following running cost problems:*

$$\inf_{A\in\mathbf{A}} \rho\left(\int_0^T r(X_{1,t}^A, A_t)dt + q(X_{1,T}^A)\right)$$

For such problems, we introduce a new state, $X_{2,t} := \int_0^t r(X_{1,s}^A, A_s)ds$, and rewrite the optimization problem as $\inf_{A\in\mathbf{A}} \rho(X_{2,T}^A + q(X)_{1,T}^A)) =: \inf_{A\in\mathbf{A}} \rho(g(X_T^A))$, where $X_t := (X_{1,t}, X_{2,t})$ and $g(x) = x_2 + q(x_1)$.

Definition 25 (Extremal Risk Measure)**.** *A function $\rho : L^2(\Omega) \to \mathbb{R}$ is said to be an extremal risk measure if there exists a convex function $f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ such that*

$$\rho(\xi) = \inf_{y\in\mathbb{R}^m} \mathbb{E}[f(\xi, y)].$$

For simplicity, we assume that f is convex with at most quadratic growth and that g is convex and Lipschitz continuous. However, we note that these growth conditions can be easily relaxed on a case-by-case basis by nothing that $X_T^A \in L^p(\Omega)$ for all $p < \infty$.

The primary motivation of this definition is the following extremal formula involving the conditional value-at-risk (CVaR) of a random variable ξ, whose distribution has no probability atom and an $\alpha \in (0,1)$:

$$CVaR_\alpha = \inf_{y\ in\mathbb{R}} \mathbb{E}\left[y + \frac{1}{1-\alpha}(\xi - y)^+\right] \tag{6.3}$$

Remark 25. *In this problem setting, if $g(X_T^A)$ has an atom, this definition can be interpreted as a lower CVaR: $CVaR^-$, cf.* [36].

The intuition behind this equality is that the optimal y is equal to VaR at probability α. Then, CVaR is equal to VaR plus the expected losses exceeding VaR divided by the probability of these losses occurring, $1-\alpha$. Miller and Yang showed that several additional risk metrics of interest in application can be written in this form. Some of these are not necessarily coherent risk measures.

Simple Conditions for Coherent Risk Measures

However, by providing simple conditions on f, we can justify the nomenclature of "extremal risk measures" under which ρ is a coherent risk measure.

Proposition 4 (Coherent (Extremal) Risk Measure)**.** *Suppose the following properties of $f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ hold:*

- *Positive-homogeneity and normalization: $f(ax, ay) = af(x,y)$ for $a > 0$ and $\inf_{y\in\mathbb{R}} f(0,y) = 0$*
- *Monotonicity: $x \mapsto f(x,y)$ is non-decreasing for each $y \in \mathbb{R}^m$*
- *Sub-additivity: $f(x_1 + x_2, y_1 + y_2) \leq f(x_1, y_1) + f(x_2, y_2)$ for $x_1, x_2 \in \mathbb{R}$ and $y_1, y_2 \in \mathbb{R}^m$*
- *Translation: For every $a \in \mathbb{R}$, there exists an invertible function $\phi : \mathbb{R}^m \to \mathbb{R}^m$ such that for every pair $(x,y) \in \mathbb{R} \times \mathbb{R}^m$ we have $f(x+a, y) = f(x, \phi(y)) + a$.*

then, the function $\rho : L^2(\Omega) \to \mathbb{R}$ *defined as*

$$\rho(\xi) := \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi, y)] \tag{6.4}$$

is a coherent (extremal) risk measure.

Proof. Recall, that the four properties of coherent risk measures are positive homogeneity, monotonicity, sub-additivity, and translation invariance. Each of these can be shown from the respective assumptions on f.

To prove positive homogeneity, we compute directly

$$\rho(a\xi) = \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(a\xi, y)] = \inf_{ay \in \mathbb{R}^m} \mathbb{E}[f(a\xi, ay)] = a \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi, y)] = a\rho(\xi)$$

for $a > 0$. In case $a = 0$, it follows that

$$\rho(0) = \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(0, y)] = 0.$$

Monotonicity: Let $\xi, \xi' \in L^2(\Omega)$ such that $\xi \le \xi'$ almost surely. Then,

$$\rho(\xi) = inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi, y)] \le \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi', y)] = \rho(\xi').$$

Sub-additivity: Fix $\xi_1, \xi_2 \in L^2(\Omega)$. We compute

$$\begin{aligned} \rho(\xi_1 + \xi_2) &= \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi_1 + \xi_2, y)] = \inf_{y_1, y_2 \in \mathbb{R}^m} \mathbb{E}[f(\xi_1 + \xi_2, y_1 + y_2)] \\ &\le \inf_{y_1, y_2 \in \mathbb{R}^m} \mathbb{E}[f(\xi_1, y_1) + f(\xi_2, y_2)] = \rho(\xi_1) + \rho(\xi_2). \end{aligned}$$

Translation-invariance: Let $\xi \in L^2(\Omega)$ and $a \in \mathbb{R}$. Then

$$\rho(\xi + a) = \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi + a, y)] = \inf_{y \in \mathbb{R}^m} \mathbb{E}[f(\xi, \phi(y))) + a] = \rho(\xi) + a.$$

□

Following Miller and Yang [36] again, we see that one important feature of the optimization problem is that, in general, it is a time-inconsistent problem. This makes it impossible to directly apply dynamic programming because there is not an analogue of the law of iterated expectations for non-Markov risk measures. To tackle this issue, one stream of research efforts focuses on developing dynamic risk measures with which sequential optimization or optimal control can be performed in a time-consistent manner. Another class of research activities acknowledges time-inconsistency as an inherent property and proposes two different solution concepts. The first approach is to reinterpret the problem as a game against ones' future self. This approach leads to a PDE system. The second method is to re-write the original problem in a form where we can apply dynamic programming in an indirect way. This approach has been used to reduce to dynamic programming in a higher-dimensional state space. The rationale in this paper is similar in spirit to this so-called indirect dynamic programming method.

However, one key advantage of the proposed approach is that the structure of the extremal risk measures allows us to perform optimization over an extra variable where the objective function can be evaluated by dynamic programming involving no additional state variables. We construct optimal time-inconsistent controls by solving an equivalent bi-level optimization problem, without lifting the state-space.

6.2. Main Results: Bi-Level Optimization and Viscosity Solutions

6.2.1. Equivalent Bi-Level Optimization

Short Summary 1 (Equivalent Bi-level Optimization)**.** *We demonstrate an equivalence between a time-inconsistent stochastic control problem involving extremal risk measures and a bi-level optimization problem in this section's theorem.*

To do so, we need to recall our goal, namely solving the generalized stochastic control problem

$$\inf_{A \in \mathbf{A}} \rho(g(X_T^A)), \tag{6.5}$$

where ρ is a fixed extremal measure and $g(X_T^A)$ is the state-dependent cost when the control A is executed. In general, we can see this problem as a time-consistent non-linear optimal stochastic control problem to which we cannot apply dynamic programming. However, Miller and Yang [36] teach us how to use the structure of ρ as an extremal risk measure to convert this into an equivalent bi-level optimization problem.

Theorem 3 (Bi-Level Optimization, cf. [36])**.** *We can write the problem of dynamic optimization over an extremal measure as*

$$\inf_{A\in\mathbf{A}} \rho(g(X_T^A)) = \inf_{y\in\mathbb{R}^m} V(y), \tag{6.6}$$

where V is defined via the standard stochastic control problem of the form

$$V(y) := \inf_{A\in\mathbb{A}} \mathbb{E}[f(g(X_T^A), y)]. \tag{6.7}$$

Proof. A complete proof can be found in Miller and Yang [36] □

Remark 26. *The value $V(y)$ depends on the initial value x_0 of X^A. For simplicity, however, we suppress the dependency and implicitly assume that we fix the initial value as x_0 for the rest of the thesis.*

At this point, we have converted the time-inconsistent stochastic control problem (6.5) to an equivalent bi-level optimization problem involving a standard stochastic control problem. For convenience, we call the right-hand sides of (6.6) and (6.7) the outer optimization problem and the inner optimization problem, respectively.

6.2.2. A Note on Assumptions

We record three main assumptions. These are used to obtain various properties of the bi-level optimization problem along the way.

Assumption 2 (Uniform Semi-Concavity)**.** *The function $y \mapsto f(x,y)$ is uniformly semi-concave for all $x \in \mathbb{R}$. That is, there exists $M > 0$ such that*

$$f(x, y+\xi) \leq f(x,y) + \xi \cdot D_y f(x,y) + \frac{1}{2}M\|\xi\|^2$$

for all $(x,y) \in \mathbb{R}\times\mathbb{R}^m$ and for all $\xi \in \mathbb{R}^m$.

Together with the convexity assumption of f, this assumption guarantees the continuous differentiability of the mapping $y \mapsto f(x,y)$ for each $x \in \mathbb{R}^m$. This regularity ultimately carries over to various smoothness results of the outer optimization problem[1] of our bi-level optimization problem. Nonetheless, Miller and Yang [36] showed, that these assumptions can be relaxed by taking inf-convolutions of f when f is not semi-concave.

Remark 27. *CVaR does not satisfy the uniform semi-concavity assumption! The relaxation is therefore particularly important from the application perspective.*

Assumption 3 (Uniform Parabolicity)**.** *There exists an $\epsilon > 0$ such that*

$$\sigma(x,a)\sigma(x,a)^{\perp} - \epsilon I_n$$

is positive semidefinite for all $(x,a) \in \mathbb{R}^n \times \mathbb{A}$.

The main use of this assumption is to construct optimal controls for the stochastic control problem (6.7). In particular, it guarantees that the viscosity solution to the Hamilton-Jacobi-Bellman equation is smooth.

Remark 28. *This constraint can be relaxed by adding additional independent Brownian motions to the dynamics of X^A.*

Assumption 4 (Convexity)**.** *The control set $\mathbb{A}$ is convex and the map $(A,y) \mapsto f(g(X_T^A),y)$ is jointly convex, almost surely.*

The primary use of this assumption is to guarantee convexity of the outer optimization problem in (6.6) and (6.6). In particular, this allows to implement a gradient descent algorithm with guaranteed convergence to a global minimum.

Record to Record Algorithm

Miller and Yang used gradient descent – a first-order optimization algorithm. In literature, it is often referred to as the steepest descent, as it is taking steps proportional to the negative of the gradient of a to be optimized function. It might however be hard to actually calculate the gradient. Approximations will therefore often be used, which

[1]Compare formula (6.6.)

increase erroneous results. Instead of Gradient Descent, one might think of using a Record-to-Record algorithm (RRT) instead.

This algorithm is also a deterministic optimization algorithm inspired by simulated annealing. The algorithm accepts a non-improving neighbor solution with objective value less than the RECORD minus a deviation D. RECORD represents the best objective value of the visited solutions during the search. The bound decreases with time as the objective value RECORD of the best found solution improves.

Data: Input initial value and $DEVIATION > 0$.
Result: Best min *Objective(RECORD)* found.
Init. $RECORD =$ initial value and $DEVIATION$
Generate a random neighbor by updating actual $RECORD$
repeat
  if Objective(Y) $< RECORD + DEVIATION$ **then**
    Accept direction
  end
  if *Objective(Y) < RECORD* **then**
    Accept neighbor solution $RECORD =$ Objective(Y)
  end
  if no improvement after k iterations, stopping criteria satisfied **then**
    Decrease $DEVIATION$
  end
until $DEVIATION$ reaches $THRESHOLD$;

Algorithm 1: Record To Record Travel

Remark 29. *Even when benchmarked to the Rosenbrock function[2], RRT finds the global minimum.*

In Figure 6.1 We see the Record-to-Record algorithm applied to the Rosenbrock function. We randomly start on the top right corner and the algorithm makes his way to its global minimum. It does not get stuck in the dark blue ring which appears to be very similar to the global minimum. This is where many algorithms have a problem when benchmarked to the Rosenbrock function. RRT however performs well.

When Assumption 3 does not hold, we can still compute so-called proximal supergradients and run a descent algorithm which converges to a local minimizer.

[2] $f(x,y) = (a-x)^2 + b(y-x^2)^2$. Further information on the Rosenbrock function can be found here: https://en.wikipedia.org/wiki/Rosenbrock_function last accessed on 30 May 2016

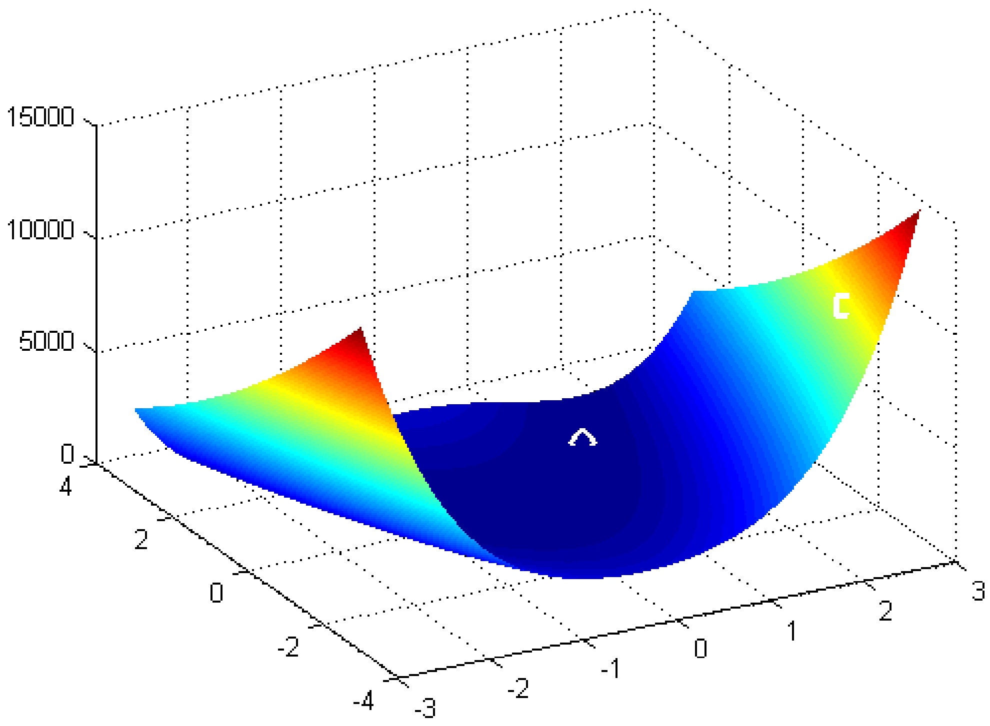

Figure 6.1.: RRT solution with the Rosenbrock function

Example 10 (The Risk Manager's Case)**.** *Along other conditions mentioned in Miller's and Yang's work, the following sufficient condition guarantees Assumption 3 to hold.*

$A \mapsto g(X_T^A)$ *is convex and* $x \mapsto f(x,y)$ *is non-decreasing convex for each* $y \in \mathbb{R}^m$. *Checking the convexity might nonetheless be a non-trivial task.*[3]

Recalling $f^{CVaR}(\xi,y) := y + 1(1-\alpha)^{-1}(\xi - y)^+, x \mapsto f^{CVaR}(x,y)$ *is non-decreasing convex and therefore, the assumption holds. Many practical examples including the mean-CVaR portfolio optimization satisfy this convexity assumption.*

6.2.3. Pseudocode

Using these results, we can solve time-inconsistent optimal control problems involving extremal risk measures via a gradient descent solution of a bilevel optimization problem.

[3] Topic of future research.

As an example of how the approximation schemes and gradient descent are utilized, consider the explicit problem:

$$\inf_{A \in \mathbf{A}} CVaR_\alpha[X_T^A].$$

We can solve this problem with the following procedure, illustrated by a pseudo-code.

6.3. Analytical Properties of the Outer Objective Function V

Let's investigate some analytical properties of the outer objective function V.

Short Summary 2. *We begin by showing the convexity of V. Then, we present a semi-concavity estimate of V at points where there exists an optimal control. Furthermore, we use this estimate and convexity to show V is differentiable at such points and to provide a probabilistic representation of the gradient.*

Theorem 4 (Convexity of V)**.** *If Assumption 3 holds, then, the outer objective function V is convex.*

Proof. Let $y, y' \in \mathbb{R}^m$ and $\theta \in [0, 1]$. For any $\epsilon > 0$, let $A, A' \in \mathbf{A}$ be ϵ-suboptimal controls such that

$$V(y) + \epsilon \geq \mathbb{E}[f(g(X_T^A), y)] \text{ and } V(y') + \epsilon \geq \mathbb{E}[f(g(X_T^{A'}), y')].$$

We can now obtain

$$V(\theta y + (1-\theta)y') \leq \mathbb{E}\left[f(g(X_T^{\theta A + (1-\theta)A'}), \theta y + (1-\theta)y')\right]$$

$$\leq \mathbb{E}\left[\theta f(g(X_T^A), y) + (1-\theta) f(g(X_T^{A'}), y')\right]$$

$$\leq \theta V(y) + (1-\theta)V(y') + \epsilon.$$

Because ϵ was arbitrary, the convexity of V follows. □

The convexity of V motivates us to use a (sub)gradient descent-type algorithm for solving the outer optimization problem. Under the semi-concavity assumption (Assumption 2), we justify gradient descent approaches by proving the differentiability of V. To that end, we record the following semi-concavity estimate.

Proposition 5. *Suppose that Assumption 2 holds. For any fixed $y \in \mathbb{R}^m$, we assume there exists $A \in \mathbf{A}$ such that*

$$V(y) = \mathbb{E}[f(g(X_T^A), y)].$$

Then, we have

$$V(y+\xi) \leq V(y) + \xi \cdot \mathbb{E}[D_y f(g(X_T^A), y)]] + \frac{1}{2} M \|\xi\|^2$$

for all $\xi \in \mathbb{R}^m$.

Proposition. Proof can be found in Miller page 9. □

Remark 30. *We assume the existence of an optimal control in this proposition. However, Miller shows that this assumption is valid in the next subsection by constructing an optimal control from an associated HJB equation under Assumption 2.*

More importantly, with Assumption 3, the function V is in fact differentiable and we have a probabilistic representation of the gradient at each point.

Theorem 5 (Differentiability)**.** *Suppose that Assumptions 1 and 3 hold. For any fixed $y \in \mathbb{R}^m$, we assume that there exists $A \in \mathbf{A}$ such that*

$$V(y) = \mathbb{E}[f(g(X_T^A), y)].$$

Then, V is differentiable at y and its gradient can be computed as

$$DV(y) = \mathbb{E}[D_y f(g(X_T^A), y)].$$

Proof. As we know from Miller and Yang's Convexity Theorem 4, V is convex and finite for all $y \in \mathbb{R}^m$. Therefore, its sub-differential is non-empty at each point. Fix $y \in \mathbb{R}^m$ and suppose $A \in \mathbf{A}$ is a control such that

$$V(y) = \mathbb{E}[f(g(X_T^A), y)].$$

Let $z \in \partial V(y)$ be an arbitrary sub-gradient of V, i.e.,

$$V(y+\xi) \geq V(y) + \xi \cdot z$$

for all $\xi \in \mathbb{R}^m$. By Proposition 5, we also have the inequality

$$V(y+\xi) \leq V(y) + \xi \cdot \mathbb{E}[D_y f(g(X_T^A), y)] + \frac{1}{2} M \|\xi\|^2$$

for all $\xi \in \mathbb{R}^m$. With these two inequalities, we can now follow

$$\xi \cdot (z - \mathbb{E}[D_y f(g(X_T^A), y)]) \leq \frac{1}{2} M \|\xi\|^2,$$

which is a contradiction – unless $z = \mathbb{E}[D_y f(g(X_T^A), y)]$. As we know that $\partial V(y)$ is single-valued, we conclude that V is differentiable at y and also that

$$DV(y) = \mathbb{E}[D_y f(g(X_T^A), y)],$$

as desired. □

With this result, we can solve the outer optimization problem using a gradient descent algorithm given that the function value $V(y)$ and its gradient $DV(y)$ are provided. We notice that $V(y)$ can be computed by dynamic programming. It is worth mentioning that the overall problem is still time-inconsistent but our bi-level decomposition allows us to solve the inner optimization problem using dynamic programming. We investigate the inner optimization problem and provide a constructive approach to solve the overall problem in the following subsection.

Partial Differential Equation Characterization of V and DV

For an initially given value x_0 of the state, we recall $V(y)$ as the optimal value of the inner optimization problem. We can solve this problem by dynamic programming. First, we compute $V(y)$ in terms of the viscosity solution of an associated HJB equation. In the process we construct an optimal control, which guarantees V is differentiable by Theorem 5. Furthermore, we can compute the gradient $DV(y)$ in terms of the viscosity solution of a linear parabolic equation.

An introduction to viscosity solutions can be found in Brokate and Witterstein [13] or Grandits [20]. By applying Brokate and Witterstein's theorem of unique viscosity solutions, we can easily conclude the following proposition:

Proposition 6. *Given $y \in \mathbb{R}^m$, let v be the viscosity solution of the HJB equation*

$$v_t + \inf_{A \in \mathbf{A}} \left[\frac{1}{2} \langle \sigma(x,A), \sigma(x,A) D_x^2 v \rangle + \mu(x,A) \cdot D_x v \right] = 0 \tag{6.8}$$

within $[0,T) \times \mathbb{R}^n$ for $v(T,x) = f(g(x),y)$ on $t = T \times \mathbb{R}^n$. Then, we have

$$V(y) = v(0, x_0),$$

where x_0 is the initial value of the SDE from Miller and Yang [36].

Due to Assumption 2, we now know that the value function v is in fact smooth by the Evans-Krylov Theorem (cf. Caffarelli et al. [14]) and associated regularity theory. This allows us to compute the gradient by solving a linear equation.

Theorem 6 (Gradient SDE)**.** *Suppose that Assumption 2-3 hold. Given $y \in \mathbb{R}^m$, let v be the viscosity solution of Proposition 6.*

1. *An optimal control, $A_t^* := a^*(t, X_t^{A^*})$, exists, where $a^* : [0,T) \times \mathbb{R}^n \mapsto \mathbb{A}$ satisfies*

$$a^*(t,x) \in \arg\min_{a \in \mathbf{A}} \left[\frac{1}{2} \langle \sigma(x,a), \sigma(x,a) D_x^2 v \rangle + \mu(x,a) \cdot D_x v \right]$$

   for all $(t,x) \in [0,T) \times \mathbb{R}^n$.

2. *Let $w : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}^m$ be the viscosity solution of the decoupled system of linear equations*

$$w_t + \frac{1}{2} \langle \sigma(x, a^*(t,x)), \sigma(x, a^*(t,x)) D_x^2 w \rangle + \mu(x, a^*(t,x)) \cdot D_x w = 0 \tag{6.9}$$

   in $[0,T) \times \mathbb{R}^n$ and

$$w(T,x) = D_y f(g(x), y)$$

   on $t = T \times \mathbb{R}^n$. Then, we have

$$DV(y) = w(0, x_0),$$

where x_0 is the initial value of the SDE.

A rigorous proof can be found in Miller and Yang [36].

Remark 31. *If w were smooth, we could apply Ito's Lemma from* [22] *for small $h > 0$ and conclude that w satisfies 6.9.However, the coefficients in (6.9) are not necessarily continuous and hence the solution w is potentially discontinuous.*

Using Proposition 6 and 6, we can calculate $V(y)$ and $DV(y)$ at each $y \in \mathbb{R}^m$ by (numerically) solving the PDEs in (6.9). Therefore, we can solve the outer optimization problem using a gradient descent-type algorithm and find a globally optimal solution due to the convexity of V.

Quick Implementation Guidelines

Once we fix a choice of λ (which controls the trade-off between CVaR and Mean) and choice of y (the parameter we optimize over, which ends up being VaR), then solving HJB can be achieved by standard finite-difference for a parabolic PDE. Attention is necessary for time-step sizing for stability because we are forced to use explicit schemes – meaning too big time steps lead to erroneous results. Solving the gradient PDE is quite tricky. In theory, the solution to this is not a smooth function, so we must use a monotone scheme. We found this to be necessary in practice as well, as alluded in Footnote 8 on Page 20 of Miller and Yang's work [36]. Further knowledge of up-winding and monotone schemes might be required.

Once we have a value and gradient for a fixed value of λ and y (from solving the PDEs above), we can just run gradient descent - in theory - assuming we have the convexity conditions outlined as in the paper. We found this part to be somewhat sensitive in practice, however. According to Miller and Yang, best practice is a bisection plus gradient descent approach. That means to optimize over y with a bisection method, but using gradient information to decide which direction to move.

Remark 32 (Quality of the Resulting Expected Shortfalls)**.** *The quality achieved from Miller and Yang's approach can be further analyzed by application of the Wasserstein metric as done by* [42].

Chapter 7.

Conclusion

This thesis broadly discusses approaches of portfolio optimization in order to better understand modern portfolio theory. We learn how to hedge earthquake risks within a Japanese stock and fixed income portfolio of an investment bank and deepen knowledge of Japanese markets. It goes beyond the Basel approaches and discusses the necessity of using more sophisticated risk measures than the Value-at-Risk.

The “Otemachi-Approach” is introduced as the process on how to decide to optimize one’s portfolio from a risk manager’s view. In testing this hypothesis, the "Otemachi Approach" revealed its potential in optimizing risk measures in modern portfolio theory. We learn which risk measures are appropriate within a fast moving environment like financial markets and it compares a variety of approaches of portfolio theory for banks to use for a regular reporting scheme.

The approach revealed that by diversifying ones portfolio by adding earthquake technology related firms (for example producers of seismic isolators) and gum as a commodity to the portfolio, we were able to hedge earthquake risks and we were able to debunk the idea of hedging by investments in the construction industry by back-testing Japanese market moves. The approach offers a step-by-step introduction to abstract a very specific problem of Japanese markets to more general risk management problems in the financial industry, i.e., portfolio optimization and risk measuring.

For deepening our understanding of modern portfolio theory, this thesis discusses examples that outline the necessity of using coherent risk measures in practice. By using an extremal representation of CVaR, we reformulate the optimal control problem as a bi-level optimization problem in which the outer optimization problem is convex and the inner optimization problem is standard stochastic optimal control. We were able to lose restrictions of Miller and Yang's model [36] by using a Record-to-Record algorithm and proposed to additionally use the Wasserstein metric as introduced by Schmock [43] to measure the quality of approximated CVaR as received from the model.

Future research topics could potentially cover the fund manager's problem for application of the bi-level optimization done by Miller and Yang [36]. Whereas classical portfolio theory studies utility maximization from the point of view of an investor, whose preferences are modeled by a concave utility function, in reality, portfolio management is commonly delegated to a fund manager. To increase the efficacy of the manager, s/he is often paid by an incentive scheme that depends on the performance of the fund s/he manages. Such a scheme can be composed, for example, of a fixed fee, some percentage of the fund, plus an additional reward, which consists of one (or a combination of several) call options on the fund. As a consequence, two differences to the classical setting arise.

- First, the utility function, under which the optimization is carried out, does not represent the preference structure of the investor (also called the principal), but rather the manager's (the agent's) preference structure.
- Second, what is optimized under this utility function is not the terminal value of the fund itself, but rather some function of it, which depends on the specific incentive scheme.

There are many approaches to optimize one's portfolio, but it is highly recommended not to solely trust in the standard Basel approaches but to back-test with other, more sophisticated methods. Especially CVaR is a proven tool to measure encountered market and credit risks in investment banking.

Bibliography

[1] Carlo Acerbi, Claudio Nordio, and Carlo Sirtori, *Expected shortfall as a tool for financial risk management*, arXiv cond-mat/0102304 (2001).

[2] Viral V Acharya, Thomas F Cooley, Matthew P Richardson, Ingo Walter, et al., *Regulating Wall Street: The Dodd-Frank Act and the new architecture of global finance* **608** (2010).

[3] Hanen Ould Ali and Faouzi Jilani, *Mean-VAR Model with Stochastic Volatility*, Procedia-Social and Behavioral Sciences **109** (2014), 558–566.

[4] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath, *Coherent measures of risk*, Mathematical finance **9** (1999), no. 3, 203–228.

[5] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, David Heath, and Hyejin Ku, *Coherent multiperiod risk adjusted values and Bellman's principle*, Annals of Operations Research **152** (2007), no. 1, 5–22.

[6] Paul J Atzberger, *An Introduction to Portfolio Theory*, http://www.math.ucsb.edu/ atzberg/finance-Major/portfolioTheory.pdf as of 23 April 2016 (2013).

[7] Jan Baldeaux, Martino Grasselli, and Eckhard Platen, *Pricing currency derivatives under the benchmark approach*, Journal of Banking & Finance **53** (2015), 34–48.

[8] Richard Bellman, *A Markovian decision process*, Journal of Mathematics and Mechanics **38** (1957), 679–684.

[9] Jonathan Benchimol and André Fourçans, *Money and risk in a DSGE framework: A Bayesian application to the Eurozone*, Journal of Macroeconomics **34** (2012), no. 1, 95–111.

[10] Dimitris Bertsimas, Geoffrey J Lauprete, and Alexander Samarov, *Shortfall as a risk measure: properties, optimization and applications*, Journal of Economic Dynamics and Control **28** (2004), no. 7, 1353–1381.

[11] Wenjie Bi, Liuqing Tian, Haiying Liu, and Xiaohong Chen, *A stochastic dynamic programming approach based on bounded rationality and application to dynamic portfolio choice*, Discrete Dynamics in Nature and Society **20** (2014).

[12] Maxim Bichuch and Stephan Sturm, *Portfolio optimization under convex incentive schemes*, Finance and Stochastics **18** (2014), no. 4, 873–915.

[13] Martin Brokate and G Witterstein, *Partielle Differentialgleichungen I*, Vorlesungsskriptum WS2002/2003, Munich University of Technology (2002).

[14] Luis Caffarelli and Luis Silvestre, *The Evans-Krylov theorem for non local fully non linear equations*, arXiv:0905.1339 (2009).

[15] Michel Crouhy, Dan Galai, and Robert Mark, *Prototype risk rating system*, Journal of banking & finance **25** (2001), no. 1, 47–95.

[16] J Daníelsson, BN Jorgensen, M Sarma, and CG de Vries, *Sub-additivity re-examined: the case for Value-at-Risk*, TU Eindhoven, 2005.

[17] Marnix Engles, *Portfolio Optimization: Beyond Markowitz*, Master's thesis, Leiden University (2004).

[18] Jun Fu, Jiaqin Wei, and Hailiang Yang, *Portfolio optimization in a regime-switching market with derivatives*, European Journal of Operational Research **233** (2014), no. 1, 184–192.

[19] Alexei A Gaivoronski and Georg Pflug, *Value-at-risk in portfolio optimization: properties and computational approach*, Journal of Risk **7** (2005), no. 2, 1–31.

[20] Peter Grandits, *Stochastische Kontrolltheorie*, Lecture Notes. FAM Deparment, Vienna University of Technology (2014).

[21] John Hull, *Risk Management and Financial Institutions*, Vol. 733, John Wiley & Sons, New York, 2012.

[22] Kiyosi Itō, *On stochastic differential equations*, Vol. 4, American Mathematical Society, 1951.

[23] Zhuo Jin, *Optimal Debt Ratio and Consumption Strategies in Financial Crisis*, Journal of Optimization Theory and Applications **166** (2015), no. 3, 1029–1050.

[24] Michael Kalkbrener, *An axiomatic approach to capital allocation*, Mathematical Finance **15** (2005), no. 3, 425–437.

[25] Hiroshi Konno and Hiroaki Yamazaki, *Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market*, Management Science **37** (1991), no. 5, 519–531.

[26] Ralf Korn and Elke Korn, *Optionsbewertung und Portfolio-Optimierung: Moderne Methoden der Finanzmathematik*, Springer-Verlag, 2013.

[27] Waikei R Lam and Kiichi Tokuoka, *Assessing the Risks to the Japanese Government Bond Market*, Journal of International Commerce, Economics and Policy **4** (2013), no. 01, 135–142.

[28] Hans-Jakob Lüthi and Jörg Doege, *Convex risk measures for portfolio optimization and concepts of flexibility*, Mathematical Programming **104** (2005), no. 2-3, 541–559.

[29] Harry M Markowitz, *Portfolio selection*, The Journal of Finance **7** (1952), no. 1, 77–91.

[30] ———, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York, 1959.

[31] ———, *Portfolio selection: efficient diversification of investments*, Yale University Press, New Haven, 1971.

[32] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts, *Quantitative risk management: Concepts, techniques and tools*, Princeton University Press, Princeton, 2015.

[33] Robert C Merton, *Lifetime portfolio selection under uncertainty: The continuous-time case*, The review of Economics and Statistics **51** (1969), no. 3, 247–257.

[34] ———, *Optimum consumption and portfolio rules in a continuous-time model*, Journal of Economic Theory **3** (1971), no. 4, 373–413.

[35] Robert C Merton et al., *An analytic derivation of the efficient portfolio frontier*, Journal of Financial and Quantitative Analysis **7** (1972), no. 4, 1851–1872.

[36] Christopher W Miller and Insoon Yang, *Optimal Control of Conditional Value-at-Risk in Continuous Time*, arXiv 1512.05015v2 (2016).

[37] Yurii Nesterov, *Smoothing technique and its applications in semidefinite optimization*, Mathematical Programming **110** (2007), no. 2, 245–259.

[38] Eckhard Platen, *A benchmark approach to finance*, Mathematical Finance **16** (2006), no. 1, 131–151.

[39] John W Pratt, *Risk aversion in the small and in the large*, Econometrica: Journal of the Econometric Society **32** (1964), no. 1/2, 122–136.

[40] András Prékopa, *Stochastic programming*, Vol. 324, Springer Science & Business Media, Amsterdam, 2013.

[41] Tyrrell R Rockafellar and Stanislav Uryasev, *Conditional value-at-risk for general loss distributions*, Journal of Banking & Finance **26** (2002), no. 7, 1443–1471.

[42] Uwe Schmock, *Modelling dependent credit risks with extensions of CreditRisk+ and application to operational risk*, Lecture Notes. FAM Department, Vienna University of Technology URL: www.fam.tuwien.ac.at/ schmock/notes/ExtentionsCreditRiskPlus.pdf (2015).

[43] ———, *Stochastic Analysis for Financial and Actuarial Mathematics*, Lecture Notes. FAM Department, Vienna University of Technology URL: www.fam.tuwien.ac.at/ schmock/notes/StochasticAnalysis_20151117.pdf (2015).

[44] Society of Actuaries, *Risk Management: The Current Financial Crisis, Lessons Learned and Future Implications*, www.soa.org (Published 2008, Last found online 23 Apr 2016).

[45] Gernot Tragler, *Nichtlineare Optimierung*, Lecture Notes. ORCOS Department, Vienna University of Technology (2014).

[46] ———, *Angewandtes Operations Research*, Lecture Notes. ORCOS Department, Vienna University of Technology (2014).

List of Figures

List of Tables

Appendix A.

Matlab Codes on Portfolio Theory

```

ticker = {'ALPMY' 'BRDCY' 'DNZOY' 'EJPRY' 'FANUY' 'KDDIY' ...
    'KMTUY' ...
     'MITSY' 'MSBHY' 'MTU' 'NSANY' 'NSSMY' 'NTT' 'PCRFY' ...
         'SFTBF' 'SHECY' ...
     'SMFG' 'SNE' 'TKOMY' 'TKPYY' 'TYIDY'}; %for quicker tests

field = 'Adj Close'; % retrieve closing price data
fromdate = '06/30/2014'; % beginning of date range for ...
    historical data
todate = '06/30/2015'; % ending of date range for historical data

for i = 1:21
     Price.(ticker{i}) = fetch(C,ticker{i},field, fromdate, ...
         todate,'w');
     temp = Price.(ticker{i});
     ClosePrice(:,i) = temp(:,2);
end

ClosingPricesTSprep = fints(temp, ClosePrice);
ClosingPricesTS = fints(ClosingPricesTSprep.dates, ...
    fts2mat(ClosingPricesTSprep),...
{'ALPMY' 'BRDCY' 'DNZOY' 'EJPRY' 'FANUY' 'KDDIY' 'KMTUY' ...
     'MITSY' 'MSBHY' 'MTU' 'NSANY' 'NSSMY' 'NTT' 'PCRFY' ...
         'SFTBF' 'SHECY' ...
     'SMFG' 'SNE' 'TKOMY' 'TKPYY' 'TYIDY'}, 1, 'Portfolio ...
         Closing Prices')
plot(ClosingPricesTS)
```

```
set(gcf, 'NextPlot', 'new');
hold;
%get returns
Returns = price2ret(ClosePrice);

ReturnsTSprep = fints(temp(2:end,:), Returns);
ReturnsTS = fints(ReturnsTSprep.dates, fts2mat(ReturnsTSprep),...
{'ALPMY' 'BRDCY' 'DNZOY' 'EJPRY' 'FANUY' 'KDDIY' 'KMTUY' ...
    'MITSY' 'MSBHY' 'MTU' 'NSANY' 'NSSMY' 'NTT' 'PCRFY' ...
        'SFTBF' 'SHECY' ...
    'SMFG' 'SNE' 'TKOMY' 'TKPYY' 'TYIDY'}, 1, 'Portfolio Returns')
plot(ReturnsTS)
set(gcf, 'NextPlot', 'new');
hold;

%define ratings as their probability of default
Rating = [0.11 0.11 0.01 0.01 0.04 0.15 0.1 0.1 0.02 0.02 ...
    0.01 0.11 0.05 0.01 0.018 0.2 0.01 0.02 0.11 0.11 0.11 ];


%plot returns of komatsu vs bridgestone on mothly basis
plot(Returns(:,1:2));
set(gcf, 'NextPlot', 'new');
ylabel(' Monthly Return ');
xlabel(' Month ');
legend('ALPMY', 'Bridgestone');
title(' ALPMY vs Bridgestone (weekly) Returns ');

lowerbound = [ 0.02 0.1 0.03 0.03 0.01 0.01 0.01 0.01 0.08 ...
   0.01 0.02 0.03 0.03 ...
    0.01 0.05 0.01 0.01 0.01 0.01 0.01 0.01 ];
upperbound = [ 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 ];
%Creating a portfolio object
portfolio = Portfolio('Name', 'JPN 21', 'NumAssets', 21, ...
   'LowerBound', ...
    lowerbound, 'Upperbound', upperbound );
portfolio = setBudget(portfolio, 1, 1); %set budget lower and ...
   upper bound
portfolio = portfolio.setAssetList(ticker);
portfolio = portfolio.setDefaultConstraints;
%disp(portfolio);

```

```
%[AssetMean, AssetCovar] = portfolio.getAssetMoments;
portfolio = portfolio.estimateAssetMoments(Returns) %expected ...
    returns

%[AssetMean,AssetCovar] = getAssetMoments(portfolio);
%portfolio = portfolio.( 'LowerBound', 0, 'Budget', 1); %no ...
    short sales

disp('Add Inequality');
%add inequality, bridgestone needs to be bigger than 10 percent
%portfolio = Portfolio('LowerBound', lowerbound, 'UpperBound', ...
    uperbound);

%disp(portfolio.NumAssets);
%disp(portfolio.AInequality);
%disp(portfolio.bInequality);

disp('Find efficient frontier');
%Find efficient frontier
%with 10 equidistant portfolios on the frontier
[prisk, preturns] = portfolio.plotFrontier(10);
[pwvt] = portfolio.estimateFrontier(10);
hold on
plot(prisk,preturns,'ored');
set(gcf, 'NextPlot', 'new');
%format_plots();

%get value at risk of the portfolio
%ValueAtRisk = portvrisk(PortReturn, PortRisk, RiskThreshold,
%PortValue)

ExpPortReturn = transpose(portfolio.AssetMean)
ExpCovariance = portfolio.AssetCovar
NumPorts = 21;
%PortWeights = [.01 .01 0.98]; %every row different weighting ...
    combination


%[PortRisk, PortReturn] = portstats(ExpPortReturn, ...
    ExpCovariance, PortWeights)

```

```
[PortRisk, PortReturn, PortWts] = ...
    portopt(ExpPortReturn,ExpCovariance, NumPorts)
set(gcf, 'NextPlot', 'new');

RisklessRate = 0.00125;
BorrowRate = 0.02;
RiskAversion = 5;

PortRisk;
PortReturn;
PortWts;
RisklessRate;
BorrowRate;
RiskAversion;


[RiskyRisk, RiskyReturn, RiskyWts, RiskyFraction, ...
     OverallRisk, OverallReturn] = portalloc(PortRisk, ...
         PortReturn, ...
     PortWts, RisklessRate, BorrowRate, RiskAversion);

portalloc(PortRisk, PortReturn, PortWts, RisklessRate, ...
    BorrowRate, RiskAversion);
set(gcf, 'NextPlot', 'new');

PortfolioBeta = 0;
for i = 1:21
StockBeta(i) = ExpCovariance(i,i)/RiskyRisk;
PortfolioBeta = PortfolioBeta + RiskyWts(1,i) * ...
    (ExpCovariance(i,i)/RiskyRisk);
end
%disp(PortfolioBeta);
%disp(StockBeta);

RiskThreshold = [0.01; 0.05; 0.10];
PortValue = 1;

disp('Value at Risk');
%maximum potential loss in the value of a portfolio over one ...
    period of time
%on a per unit basis
```

```
ValueAtRisk = portvrisk(RiskyReturn, RiskyRisk, RiskThreshold, ...
    PortValue)

%with real numbers: with a chance of 99,95,90 percent, the ...
    maximum loss is...
ValueAtRisk = portvrisk(RiskyReturn, RiskyRisk, RiskThreshold, ...
    10^6)



%the conditional VaR is also known as expected shortfall

%Cvar takes a look at the end of distribution of losses in ...
    scenarios
%get portfolio and calculate the losses for several different ...
    scenarios

%mean-absolute deviation
%konno yamazaki



%Monte Carlo Simulation
ReturnSeries = portsim(ExpPortReturn, ExpCovariance, 1, 1, 100);
%histogram(ReturnSeries)

%VaR mit Monte Carlo?
for i=1:21
    for j=1:100
        MCReturnsAfter10Days(j,i) = ReturnSeries(1,i,j);
    end
end
MCReturnsAfter10DaysCopy = MCReturnsAfter10Days;
%get loss distribution, with positive losses
MCReturnsAfter10Days(MCReturnsAfter10Days>0)=0;
MCLosses = MCReturnsAfter10Days*(-1);
histogram(MCLosses);
set(gcf, 'NextPlot', 'new');
ylabel(' # of portfolio returns ');
xlabel(' loss ');
title(' loss distribution ');

```

```
%pmc = PortfolioCVaR;
pmc = PortfolioCVaR('Name', 'JPN 21 CVaR', 'NumAssets', 21, ...
    'LowerBound', ...
    lowerbound, 'Upperbound', upperbound );
pmc = setBudget(pmc, 1, 1);
pmc = pmc.setAssetList(ticker);
pmc = pmc.setScenarios(MCReturnsAfter10DaysCopy);
pmc = pmc.setDefaultConstraints;
pmc = pmc.setProbabilityLevel(0.95);

figure; [pmcRisk, pmcReturns] = pmc.plotFrontier(10);
set(gcf, 'NextPlot', 'new');

%now compare Mean-Varianc and CVaR
%to really compare it, we need to transform one risk measure ...
    into another
%therefore, we now calculate the mean variance risk of the ...
    cvar portfolio

pmcwgts = pmc.estimateFrontier(10);
pmcRiskStd = pmc.estimatePortStd(pmcwgts);

%Now add CVaR portfolios to Mean-Variance plot
figure;
portfolio.plotFrontier(10);
hold on
plot(pmcRiskStd,pmcReturns,'-r', 'LineWidth', 2);
legend('Mean-Variance Efficient Frontier',...
    'CVaR Efficient Frontier', ...
    'Location','SouthEast')

set(gcf, 'NextPlot', 'new');

%compare weights
figure;
subplot(1,2,1);
area(pmcwgts');
title('CVaR Portfolio Weights');

subplot(1,2,2)
area(pwvt');
title(' Mean Variance Portfolio Weights ');
```

```

set(get(gcf,'Children'),'YLim',[0 1]);
legend(portfolio.AssetList);
set(gcf, 'NextPlot', 'new');

%we can see quite a huge difference. a possible explanation ...
    is, that for
%the mean variance model, we use the assumption of normal ...
    distribution
%which is clearly not the case.

figure;
hist(Returns(:),21);
histfit(Returns(:),21); figure(gcf);
set(gcf, 'NextPlot', 'new');


%compare to portfolio absolute mean deviation model
pmad = PortfolioMAD('Name', 'JPN 21 CVaR', 'NumAssets', 21, ...
    'LowerBound', ...
     lowerbound, 'Upperbound', upperbound );
pmad = setBudget(pmad, 1, 1);
pmad = pmad.setAssetList(ticker);
pmad = pmad.setScenarios(MCReturnsAfter10DaysCopy);
pmad = pmad.setDefaultConstraints;

figure; [pmadRisk, pmadReturns] = pmad.plotFrontier(10);
set(gcf, 'NextPlot', 'new');

pmadwgts = pmad.estimateFrontier(10);
pmadRiskStd = pmad.estimatePortStd(pmadwgts);

%figure; [pmadRisk, pmadReturns] = pmad.plotFrontier(10);
%set(gcf, 'NextPlot', 'new');

figure;
portfolio.plotFrontier(10);
hold on
plot(pmadRiskStd,pmadReturns,'g', 'LineWidth', 2);
hold on
plot(pmcRiskStd,pmcReturns,'-r', 'LineWidth', 2);
legend('Mean-Variance Efficient Frontier',...
```

```
    'MAD Efficient Frontier',...
    'CVaR Efficient Frontier', ...
    'Location','SouthEast')
set(gcf, 'NextPlot', 'new');

%compare weights
figure;
subplot(1,3,1);
area(pmcwgts');
title('CVaR Portfolio Weights');
subplot(1,3,2)
area(pwvt');
title(' Mean Variance Portfolio Weights ');
subplot(1,3,3)
area(pmadwgts');
title(' MAD Portfolio Weights ');
set(get(gcf,'Children'),'YLim',[0 1]);
legend(portfolio.AssetList);
set(gcf, 'NextPlot', 'new');
```